

Open Research Online

The Open University's repository of research publications and other research outputs

The assessment of 'wicked' competences

Other

How to cite:

Knight, Peter and Page, Anna (2007). The assessment of 'wicked' competences. Practice-based Professional Learning Centre, Centre for Excellence in Teaching and Learning, The Open University.

For guidance on citations see [FAQs](#).

© 2007 Peter Knight; 2007 Anna Page; 2007 The Open University



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

The assessment of ‘wicked’ competences

Report to the Practice-based Professional Learning Centre

“Some problems are so complex that you have to be highly intelligent and well informed just to be undecided about them.”

Peter Knight and Anna Page: peter.knight@open.ac.uk and a.c.page@open.ac.uk

In a nutshell

Empirical and other enquiries indicate that the assessment of ‘wicked’ competences is problematic. This is unfortunate since ‘wicked’ competences include ‘soft skills’ and other attitudes, skills and dispositions that are highly valued by employers.

Empirical evidence that there are, in fact, few problems in the assessment of ‘wicked’ competences may indicate that there is a problem of ‘false consciousness’, with practitioners simply not seeing how limited their assessment plans are. If that is the case, then not only is there a problem of improving assessment, there is also a far bigger problem in helping colleagues to appreciate that current practices may have pragmatic value but have very limited value as scientific descriptions of achievement and competence.

Colleagues are invited to join in further research and design activities by contacting Anna Page in the first instance

Contents

Acknowledgements.....	1
Overview and summary.....	2
Chapter 1 Assessing human achievements	4
Method in social science	5
'Normal' social science	5
Social science and mess	5
Issues in assessing student learning.....	9
Chapter 2 The assessment of 'wicked' competences	11
1. The issue	11
'Wicked' competences	11
'Wicked' competences and the like	12
Fair assessment.....	13
Tests of 'wicked' competences.....	13
2. The design of a study of the assessment of wicked competences.....	16
3. Nine wicked competences	17
The nine	17
Will clear criteria deal with assessment problems?	17
4. The six subject areas: wicked competences and their assessment	19
Accounting	19
Early years education	20
Nursing.....	21
Secondary school education	22
Social work.....	23
Youth work.....	25
5. The competences	27
a. Developing supportive relationships.....	27
b. Emotional intelligence (EI).....	27
c. Group work	28
d. Listening and assimilation	30
e. Oral communication	31
f. Professional subject knowledge.....	32
g. Relating to clients.....	33
h. Self-management.....	33
i. 'Taking it onwards' – acting on diagnoses.....	33
6. Reprise: issues in the assessment of wicked competences.....	34
1. Metaphysics.....	34
2. Fuzzy achievements	34
3. Cost and convenience	35
4. Reliability	35
5. Validity	35
6. Test standardisation.....	36
7. Performances and proxies	36
8. Application of criteria.	37
9. Direct and implied judgements.....	37
10. Assessment as an appropriate focus	37
Issues and competences	37
Table 6.1. Issues and competences	38
Glossary for Table 6.1	38
7. Survey data	39
a. Developing supportive relationships	41
Difficulty.....	41
Assessment methods in use.....	41
b. Emotional intelligence.....	42
Difficulty.....	42
Assessment methods in use.....	42

c. Group work	43
Difficulty	43
Assessment methods in use	43
d. Listening and assimilating	44
Difficulty	44
Assessment methods in use	44
e. Oral communication	45
Difficulty	45
Assessment methods in use	45
f. Professional subject knowledge	46
Difficulty	46
Assessment methods in use	46
g. Relating to clients	47
Difficulty	47
Assessment methods in use	47
h. Self-management (confidence and effectiveness)	48
Difficulty	48
Assessment methods in use	48
i. 'Taking it onwards' – acting on diagnoses	49
Difficulty	49
Assessment methods in use	49
Discussion	50
8. Telephone interviews	51
Analysis of Accounting interviews	52
Analysis of Nursing interviews	55
Analysis of Social Work interviews	57
The three hypotheses	59
There is no great problem assessing 'wicked' competences	59
Unexpected survey results are attributable to the use of inappropriate enquiry methods	59
Unexpected survey responses are attributable to teachers' 'false consciousness'	59
Other interview findings	59
9. A summary of the empirical enquiry into the assessment of 'wicked' competences	61
Chapter 3 A pragmatic approach to the assessment of 'wicked' competences	62
Further investigations	62
Suggestions for improving assessment practices	64
Teachers	65
Departments	67
Institutions	69
Students	70
References	72
Appendix 1: Extracts from the Wikipedia entry 'wicked problem'	75
Appendix 2: Extracts from pbpl document <i>Developing systems to support pbl</i> , January 2007	76

Acknowledgements

This work is sponsored by the Open University's Practice-based Professional Learning Centre, which is one of the University's four Centres for Excellence in Teaching and Learning.

The quotation on the cover is attributed to Laurence J. Peter by Conklin (2003:1).

Overview and summary

"Wicked" competences are achievements that cannot be neatly pre-specified, take time to develop and resist measurement-based approaches to assessment. On the basis of knowledge of assessment practices in higher education in general, it was anticipated that there would be acute problems assessing this class of outcomes. And they are also important outcomes of *higher* education, since they are widely valued by employers and smooth the path of study and other forms of research¹.

This study identified nine such competences in a total of six subject areas. Data were collected in the autumn of 2006 from 83 respondents on the experience of assessing eight of the competences. Findings were organised around three themes: the difficulties of assessing them; the priority given to assessing them in higher education; and means by which they were assessed.

There are no particular surprises in the means used to assess these competences; nor was it surprising to see that they did not always get the priority that might have been expected.

The surprising finding was that they were not seen to be especially difficult to assess. Three main lines of explanation for this are considered: one explanation is that 'wicked' competences are not really hard to assess; a second is that enquiry methods were not fit for the purpose; the third is that there is a degree of false consciousness amongst respondents.

Follow-up telephone interview with 14 respondents did disclose a lot of problems in the assessment of 'wicked' competences. Interview data favour the second and third explanations. However, they do not allow us to prefer one over the other. Indeed, it is suggested that both explanations may be simultaneously sustained and that it may be provisionally assumed that the assessment of 'wicked' competences is indeed problematic.

This suite of enquiries, empirical and otherwise, prompts four recommendations.

1. Further enquiries could be justified, although close-up research is inherently expensive.
2. Any interventions to enhance the assessment of 'wicked' competences should begin by helping colleagues to appreciate the inadequacies of current practices that are typically – and wrongly – assumed to be 'good enough'.
3. Higher education institutions might usefully consider the four sets of proposals for enhancing assessment practice in general that are set out in part 3 of this report. Doing so should attenuate some problems attaching to the assessment of 'wicked' competences. This will be particularly useful for those concerned with the design and enactment of work- or practice-based professional courses.
4. There is good reason to suppose that the formation and assessment of 'wicked' competences should best be conceived in terms of programme², not course design. Admittedly, there is little research into programme design and virtually none on programme assessment. However, work (Huber *et al.*, 2007) in the USA on integrative learning seems to be tracking similar territory to this enquiry into 'wicked' competences, with two crucial differences. First this work also attends to means of fostering such competences and secondly, it is adamant that programme approaches are necessary

¹ To be scrupulous, there is some evidence that a lack of 'wicked' competences, employability skills and the like on graduation may not matter after three years of graduate employment. Mason and colleagues (2003) found some indications that differences become evened out through (probably non-formal) learning in the workplace. There is a catch though. Those with employability skills or 'wicked' competences were more likely to be employed in the first place. If that is the case, institutions concerned with employability will be equally concerned that their graduates are engaged in ways that foster the competences that get them on the job ladder.

² A programme is the set of units, courses or modules leading to an award, such as a Diploma, bachelor's or doctor's degree.

We invite colleagues in professional subject areas to consider undertaking 'close-up' work to get further purchase on this issue, perhaps with a view of offering a symposium to the *Higher Education Close-up* conference in Cape Town, 26-28 June 2008.

We also invite colleagues to consider forming an inter-disciplinary network to consider designs that favour the development of 'wicked' competences, integrative learning and 'soft skills' in general.

Chapter 1 Assessing human achievements

Arts, humanities and social sciences are concerned with the distinctively *human* world. It follows that the assessment of human achievements should take its methods from those domains, rather than from natural science.

Arts and humanities, with their emphasis on the particular, would supply methods, especially heuristic methods, for assessing the quality of individual achievements. Eisner (1985) has drawn on these methods to propose that *connoisseurship* is an appropriate way of evaluating educational phenomena.

Without denying the significance of the particular, the use of the term 'sciences' in 'social sciences' marks out these subjects' intention of creating some forms of generalisation. Similarly, assessment practice is often concerned to create generalisations, particularly generalisations from now to the future, as in statements of ability and competence. It also often involves making comparative judgements about groups, often large groups, of people, and thereby supports the making of inferences about individuals on the basis of data about their performance in general assessments of achievement. As such, and without prejudice to the connoisseurship approach derived from arts and humanities, assessment practice can be identified with the concerns and practices of the social sciences.

What are those practices?

Method in social science

'Normal' social science

Many social phenomena can be treated, for many intents and purposes, as if they are objective. Patterns of movement, incidence of crime, voting preferences, expenditure – these are all phenomena that can be treated as if they were objective, rather like the facts of natural science. As such, they may also be subject to methods akin to those of the natural sciences.

This might be described as 'normal' social science: phenomena are treated as tolerably-determinate and are described by using quasi-scientific methods to ensure that neutral enquiries produce secure statements about the phenomena in question.

This is important social science and there are clear, stringent rules of procedure covering definition, validity, reliability, sampling, inference and reporting.

If the assessment of students' learning is seen as 'normal' social science, where the object is to establish with tolerable certainty whether they do know certain information, can apply certain procedures and perform certain tasks, then there is a very clear set of procedures that, according to 'normal' social science, ought to be used. To be more precise, they ought to be used if the intention is to produce statements that will have some high-stakes value as robust statements of the facts of student performance. The standards of 'normal' social science may be waived if the intention is simply to produce impressions that stimulate learning conversations.

Nevertheless, the high stakes assessment practices in universities seem to fall far short of the standards of 'normal' social science or, more modestly, of public examination boards. To that degree, limited confidence can be put in high stakes judgements once they are extracted from the context in which they were created.

However, these strictures cannot apply to the assessment of 'wicked' competences because they cannot be 'treated as if they were objective, rather like the facts of natural science'. This claim is defended in the next chapter. Assuming that the claim is defensible, do researchers in social science have any theories that would help us to think better about how these competences might be assessed?

Social science and mess

One account of the methods of social science enquiry, (Knight, 2001), emphasises the degree to which they can be enquiries into human conditions of uncertainty, contingency and complexity, concerns also elaborated by, for example, van Geert, 1994; Bryne, 1998; Thrift, 1999). Law (2004) has similar concerns, which he approaches in the light of actor-network thinking. The book opens with a picture of a mess of jottings, lines and clippings. Underneath it says 'If this is an awful mess ... then would something less messy make a mess of describing it' (p. 1, ellipsis in original). He continues:

This book is about that caption, and about what happens when social science tries to describe something complex, diffuse and messy. The answer, I will argue, is that it needs to make a mess of it. This is because simple clear descriptions don't work if what they are describing is not itself very coherent. The very attempt to be clear simply increases the mess. (p. 2)

He recognises the need to differentiate, much as I have distinguished between the assessment of tolerably-determinate achievements and complex ones (or 'wicked competences')

No doubt some things in the world can indeed be made clear and definite ... but alongside such phenomena the world is also textured in quite different ways. My argument is that academic methods of inquiry don't really catch these. (p. 2)

Consequently,

It cannot be the case, then, that standard research methods are straightforwardly wrong. They are significant and will properly remain so. This is why I say that I am after a broader and more generous sense of method ... I want to argue that while standard methods are very good at what they do, they are badly adapted to the study of the ephemeral, the infinite and the irregular ... the problem is not so much the standard research methods themselves, but the normativities that are attached to them in discourses about method. If 'research' methods are allowed to claim methodological hegemony or (even worse) monopoly ..., then when we are put into relation with such methods we are being placed, however rebelliously, in a set of constraining normative blinkers (p. 4)

This is an important point because there is no intention of denying that there is a swathe of graduate achievements that can in some measure be appropriately treated as determinate objects. They *can* be described by classic assessment methods. However, there are others that cannot – indeed, should not – be treated thus. In Law's words, this involves moving

... from the moralist idea that if only you do your methods correctly you will lead a healthy research life – the idea that you will discover singular truths about which all reasonable people can at least temporarily agree ... I want to subvert method by ... escaping the postulate of singularity and, and responding creatively to a world that is taken to be composed of an excess of generative forces and relations.

To do this we need to unmake many of our methodological habits, including the desire for certainty; the expectation that we can usually arrive at more or less stable conclusions about the way things really are; the belief that as social scientists we have special insights that allow us to see further than others into certain parts of social reality; and the expectations of generality that are often wrapped up in what is often called 'universalism'. (p. 9).

The claim that old methodological habits are not as robust as might be assumed depends on calling into question the metaphysical taken-for-granted that underpin those habits.

... I want to tease out some of the metaphysical assumptions that Euro-American people tend to ... think ... about reality, about what is, about ontology. First, and most generally, it appears that our experience is widely if not universally built around a sense that there is, indeed, a reality that is out there beyond ourselves '*out-there-ness*' Most of us, I guess, implicitly commit ourselves to the further sense that this external reality is usually independent of our actions and especially of our perceptions ... I call it a commitment to '*independence*'. Another more or less related common sense is that this external reality comes before us, that it precedes us ... '*anteriority*'. A further common-sense is that external reality has, or is composed of, a set of definite forms or relations ... '*definiteness*'. Another common-sense is that the world is shared, common, the same everywhere ... '*singularity*'. (pp. 24/5)

To repeat the point, there are achievements that can be fairly treated as if they did fit snugly with such assumptions. However, as we shall see in the analysis of the nine wicked competences addressed in the empirical study, it is not at all obvious that they really do fit these metaphysics.

Law develops his metaphysical critique with reference to a set of empirical studies that enacted social practices as assemblages of different metaphysics. First, a study of the medical phenomenon of arteriosclerosis:

... the empirical and matter-of-fact answers about claudification [a medical process] travel so far, but only so far. The same is the case for any other argument about out-there-ness. How far does it carry? So far, but only so far ... It is possible to observe, in one way very matter-of-factly, that the world, its knowledges, and the various senses of what is right and just, overlap and shade off into one another ... it appears, then, that in practice there are plenty of partial connections, partial inclusions, partial relations. It also appears that these do not reduce to one another. (pp. 63/4)

Second, a study of alcoholic liver disease:

our own object of study and its contexts were constantly moving about. Thus we found that we not only were shifting between different alcoholic liver diseases but also, and uneasily, between different problems. Initially we were invited to explore the diagnosis and treatment of alcoholic liver disease. Call this object number one. But as we moved into the study and interviewed the professionals, we found that we were often talking about liver disease (object number two), rather than alcoholic liver disease. Or, more specifically, we were discussing alcoholic cirrhosis (number three). Or, very commonly, the talk was of alcohol abuse and its implications for individuals and the health care system (four) Or (not necessarily the same thing) it was of alcoholism (five) Or ... it might be about overall quality of life in relation to substance abuse (six). The issue, then, was how to think about this displacement: the fact that the object of study seemed to slip and slide from one interview to the next We gradually came to think that this was not simply a sign of shoddy method, a failure to get to grips with something definite. Instead, we slowly came to believe that we were dealing with an object that wasn't fixed, an object that moved and slipped between different practices in different sites. This was an object that, as it moved and slipped, also changed its shape. It was a shape-changing object that, even more misleadingly, also changed its name. (pp. 78/9)

In these discussions of empirical work he is producing depictions of overlap, interference, miasma and non-coherence. His account is interesting because it is quite easy to imagine how assessment could be represented as a fluid and non-coherent phenomenon, like alcoholic liver disease. As another of Law's papers (2000) argues, the object – whether alcoholic liver disease or assessment – does not necessarily present itself as a well-defined and stable thing (see also Checkland, 1981).

So, rather than recommending that superior methods would lead to the creation of coherence and, with it, of general meaning, Law argues that:

Non-coherence is not necessarily a good ... But neither is coherence necessarily a good ... However, the problem with Euro-American metaphysics is its lack of symmetry. It simply assumes that that coherence is good, and tries to enact it into being. It makes no space for the acknowledgement of non-coherence. (p. 98)

One, but by no means the only, problem with trying to 'pin down' the object of enquiry, is the amount of data available to us. There is too much. In assessment, a common response to unsatisfying practices has often been to increase the amount of information that we have about student achievement, to engage in the process of exquisite description (Knight, 2006a). Portfolios, records of achievement and some transcripts are examples of depiction in ever-increasing detail. Let us leave aside the 'Shandy paradox' that we can never describe fully enough, and listen, instead, to Law's view that sensemaking, of which assessment is a particular form, involves excluding – that inclusion also presumes exclusion. Law talks about his ethnographic research and the experience of being overwhelmed by data .

I tended to think that this was my own particular problem: that I wasn't coping properly with the incessant demands of ethnography. I wondered if a better ethnographer would have been on top of all the detail and better able to keep track of the ethnographic equivalent of boxes of wood screws. However, I now think that something much more interesting and important was going on. It was that in the ethnographic method assemblage the practices I needed to make certain silences and unrealities were not in place. I was being overwhelmed by too many inscriptions of traces-in-here and the manifestations of too many realities out-there. (p. 108)

This gives principled grounds for resisting the view that better assessment involves richer descriptions of achievement and context. proliferation of data – proliferation is metaphysically naïve. Instead, Law's metaphysics lead to a view in which:

There is no general world and there are no general rules. Instead there are only specific and enacted overlaps between provisionally congealed realities that have to be crafted in a way that responds to and produces particular versions of the good that can only ever travel so far. The general, then, disappears, along with the universal. The idea of the universal transportability of universal knowledge was always a chimera ... Instead we are left with situated enactments and sets of partial connections, and it to those that we owe our heterogeneous responsibilities (p. 155)

If this is so, then the quest for assessment practices that yield general statements about 'wicked' competences and complex achievements is metaphysically hopeless – and in that case there is no rescuing the fact and Knight's (2006b) position that assessment practices only generate statements with local meanings, holds good. But, that does not mean that it is therefore impossible to imagine fertile assessment practices that engage with the complex achievements that should characterise higher education. All it says is that certain generalisations are not feasible. Indeed, given that summative assessment and warranting will not creep away, it is necessary to take a pragmatic turn and imagine ways in which judgements of some value might be made and communicated.

That imagining, though, is best based upon a tolerably-secure appreciation of existing practices in assessment. Given that so much work has been done on the assessment of more-or-less determinate achievements, the need is for a better understanding of assessment practices associated with less determinate achievements or 'wicked' competences.

The argument has been that assessment practices in higher education are deficient. When it comes to assessing tolerably-determinate achievements, university assessment practices seldom have the rigour that is expected by 'normal' social science. When it comes to the less determinate, or 'wicked', outcomes, work such as Law's sets a challenge which has hardly been acknowledged in higher education. This argument is consolidated with a closing note that summarises some of the disarray in assessment practices in universities and colleges.

Issues in assessing student learning

A convenient feature of Law's account of method in social science is that it allows us to see commonplace assessment problems as manifestations of the greater problem of knowing about social life. This works two ways. On the one hand it rescues us from assumptions that assessment problems arise from incompetence, lack of commitment or lack of investment. On the other, it casts assessment as a social practice and invites assessment practitioners to align and learn with others trying to describe social practices.

The assessment of student learning appears to be a universal social practice in which enormous amounts of time, money, ingenuity (both compliant and subversive), care and identity are invested.

It is also a recurrent subject of various sorts of concern. For example:

1. Internationally, there is a continuing quest for more (a) reliable (trustworthy), (b) valid (properly representative of the achievement being assessed), (c) useful and (d) affordable practices.
2. Success is elusive, despite the promise afforded by new technologies, which certainly allow for slicker judgements of determinate achievements.
3. Improvements on one of the four counts tends to be offset by deteriorations in one or more of the others. For example, reliability tends to compromise validity, validity to compromise economy.
4. Scholars of assessment practice seldom connect the business of assessing students achievements with literatures on social science research methods, despite the essential sameness of the ventures.
5. Consequently, assessment practice is seldom considered to raise ontological and epistemological issues, although they would normally be considered in the design of social science enquiries (Knight, 2001).
6. Assessment takes place for a variety of reasons including (a) to give students feedback on achievement in order to shape future learning (b) to describe student achievement to outsiders, especially employers and graduate schools (c) to direct students' attention to the most salient parts of a curriculum (d) to provide data that can be used for management purposes. Attention to one of the four may compromise others.
7. There is a generous range of assessment methods in currency in higher education. Methods are not tied to any particular assessment purpose, although some may be more congenial to some purposes than others.
8. What often present to individual teachers as assessment problems are often curriculum design or programme management problems.
9. It is recognised, most recently by Boud and Falchikov (2006) and Knight (2006) that assessment and curriculum practices are woven together. Although it has been a shibboleth that if the assessment is right then good learning will follow, there is an alternative view that the design of learning environments, which, understood generously, might be understood as the design of curricula, should take priority.
10. Assessment is mainly a local craft activity, taking place in single modules, units or courses. This militates against programmatic (or comprehensive) assessments of achievements that are slow-growing (Claxton, 1998) and which are outcomes of years of study, not of a few weeks spent on a single course.

These points will support the conclusion that assessment is not a singular practice but a family of diverse practices. It is not a hopeless multitude of practices but nor is it a unity. Consequently, rather than seeking generalised or all-purpose assessment solutions, we will have to settle for plurality and contingency.

Let us agree that there is a whole range of learning achievements that can be tolerably well described in advance and which can be readily assessed with high degrees of agreement. These achievements may be called 'determinate' or 'quasi-determinate'. They include knowledge of information; application of formulae and procedures in defined and straightforward situations; and the application of other well-tried heuristics. The claim, established in preceding sections, is that where the stakes are high and the intention is to produce a warrant or public attestation to achievement, then the rigour of 'normal' social science should be used.

But what of the assessment of 'wicked' competences, of achievements that are not determinate? The following study of 'wicked' competences moves assessment discourse away from the tame questions of how best to assess knowledge of Portuguese irregular verbs, the procedures for volumetric analysis of chemical compounds or differentiation in basic calculus. It asks about the assessment of non-determinate and subjective achievements.

The third chapter takes a pragmatic turn. It assumes that assessment is a social ritual that will not be banished by scholarly chants and talismans. It enquires about ways of recognising the radical uncertainty and contingency immanent in the first two points while also appreciating that assessment, like the juggernaut, will roll on.

Chapter 2 The assessment of 'wicked' competences

1. The issue

'Wicked' competence are ones that are hard to define and that cause assessors lots of problems. Yet these competences are often soft skills and other complex achievements that graduate employers say they value. The term echoes the notion of 'wicked' problems (Conklin, 2003) and, to a lesser degree, 'troublesome knowledge' (Perkins, 2006)

Can universities' claims to foster these competences be accepted, given the general view that what is not assessed is not taken seriously in the curriculum? A response to the larger question hinges on answering others: How do universities assess wicked competences? Can we trust their assessments? How do they communicate achievement in respect of these competences? And, do the assessment practices raise questions about the teaching of these wicked competences?

'Wicked' competences

The term 'wicked' competences is inspired by the idea that there are 'wicked' problems, which is elaborated in Appendix 1. These are the problems that can occupy unlimited time, arouse passions and which are ultimately beyond consensual resolution. The allocation of car parking space on university campus is a good example of a 'wicked' problem. Conklin (2003) says the four defining features of a wicked problem are that:

1. The problem is not understood until after formulation of a solution
2. Stakeholders have radically different world views and different frames for understanding the problem.
3. Constraints and resources to solve the problem change over time.
4. The problem is never solved.

They are also 'wicked' because they appear to be soluble by the application of evidence and reason and they suck people in to trying to solve rationally issues that only admit of provisional solutions, perhaps arrived at by the exercise of power.

Many assume that difficulties assessing 'wicked' competences similarly stem from shortfalls of evidence or reason. This is a category mistake. Correctly understood, the problem is that these 'competences' are usually social practices that are highly context-dependent. Judgements of performance cannot ignore the context, which makes generalisation about performance – which is, after all, what summative assessment³ creates – very difficult, if not impossible. In the order of things, 'wicked competences' may not be amenable to assessment as it is conventionally understood. This is not a radical claim that they lay beyond judgement but a more precise claim that:

1. The competences themselves are a mix of dispositions, understandings, attributes and practices.
2. They are typically non-determinate, in the sense that it is seldom possible to specify fully what it would mean to be competent in, say, emotional intelligence.
3. They take time to form, being the product of years, rather than of weeks.
4. Descriptions of performance have to include descriptions of the context, which contrasts with, for example, descriptions of IQ.
5. To understand the meaning of careful descriptions and judgements of performance we need to understand criteria and the conditions under which the performance was achieved.

³ Much of this paper concerns summative assessment, which is assessment designed to produce judgements of achievement that are robust and public – assessments that sum up achievement and that count towards awards, grades and classifications. Formative assessment will use the same techniques but with the aim of creating feedback to help students and staff improve future learning.

6. Assessments will often be expressed on a nominal or ordinal scale. If numbers are used, they should be understood as 'tags' for descriptions or for rank order in a group. They are not parametric and should not be subject to parametric manipulations.

'Wicked' competences and the like

The set of understandings, dispositions, practices and qualities that could be described as 'wicked competences' can be described by other terms. Barrie's (2004) account of the Australian concept of 'graduate achievements' includes many wicked competences. Research into what employers look for in new graduate hires does, as the following three examples show.

1. Lee Harvey and colleagues (1997) found that employers want graduates with knowledge; intellect; willingness to learn; self-management skills; communication skills; team-working; interpersonal skills.
2. Yorke (1999) found that small enterprises especially valued skill at oral communication, handling one's own work load, team-working, managing others, getting to the heart of problems, critical analysis, summarising, and group problem-solving. Valued attributes included being able to work under pressure, commitment, working varied hours, dependability, imagination/creativity, getting on with people, and willingness to learn.
3. Brennan and colleagues (2001) highlighted the significance of: initiative; working independently; working under pressure; oral communication skills; accuracy, attention to detail; time management; adaptability; working in a team; taking responsibility and decisions; planning coordinating and organizing.

The EC's expert group looking at higher education/research relations in the European Research Area (Knight, 2003) proposed that researchers should have a set of core competences, most of which could be reckoned as wicked competences:

- Reasoning, appropriate to the topic and tradition of enquiry; critical thinking, problem-setting and problem-solving of various kinds.
- Creativity/curiosity.
- Team-working and collaboration.
- Information handling, including interpretation and evaluation.
- Working across subject boundaries – multi-, inter- and trans-disciplinary practices.
- Managing projects and leading: being entrepreneurial, whether in self-employment or when working for with-profit or non-profit enterprises.
- Ethical practice.
- Dissemination, including writing for various audiences, making presentations, participating in conferences, fora and panels.

They can also be seen as a sub-set of higher education's work to foster citizenship (Colby and colleagues, 2003) and resemble the outcomes of Cardinal Newman's view of higher education and are widely valued by academic staff.

Coming at things from a different direction, Gardner (2007) identifies 'five minds for the future', which comprise⁴:

- The disciplined mind has mastered one or more professions, arts, crafts or disciplines
- The synthesizing mind determines what is important and what is fluff
- The creating mind thinks outside the box

⁴ These quotations are taken from a handout at Gardner's OU lecture 9 October 2006.

- The respectful mind ... a respectful person seeks to understand others, to work with them, to instill a respectful environment at work and in the community.
- The ethical mind requires an abstract attitude. The individual develops a concept of him or herself as a person, a worker, and a citizen

If this analysis is as persuasive as his earlier (1983) concept of 'intelligences', then there will be a wish to judge people in terms of them. However, these are non-determinate concepts – 'wicked' frames of mind.

Summarising these positions, it appears that:

- Similar-seeming things tend to attract a range of different names
- There is a lot of variation in the things that could be described as 'wicked' competences
- These lists include a mix of qualities, understandings, dispositions and practices. Some might fairly be called competences but by no means all – or even most – of them.

Fair assessment

As to the assessment of such achievements, there are two main schools of thought. One says that they must be assessed – by which its adherents seem to mean 'summatively assessed'. This view itself broadly divides into those who think that they can be measured (for a critical thinking example, see <http://www.insightassessment.com/test-cctst.html>) and those who think they can be appreciatively described. Although there is no shortage of advice about good assessment practices, it does not really solve the problem of assessing hard-to-define or 'fuzzy' outcomes (Shay, 2005; Yorke, 2005a). For example, some advice assumes observations of *in situ* activity, which, in the interests of maximizing reliability, need to be repeated over time in different situations, all of which must allow displays of achievement on the targeted competences. This approach has logistic and practical problems. An alternative is to rely on proxies for direct observation. These can include OSCEs (objective, structured clinical examinations), simulations, assessment centres, problem-working activities and portfolios. Some of these approaches are less authentic than others (and therefore less likely to be good predictors of performance) and some can be quite expensive (assessment centres). There is agreement on one thing, namely that good criteria are needed, even though we know (a) it is challenging to imagine 'good' criteria when it comes to fuzzy outcomes (b) that criteria are re-interpreted in different ways by different assessors (Price and Rust, 1999). In other words, best advice is either unrealistic (on cost and logistic grounds) or vitiated by conceptual issues (the creation and efficacy of assessment criteria)

The other response to the assessment problem is to place these achievements outside the assessment system – to make them non-assessed. There are four disadvantages:

1. Mention has already been made of the view that what is not assessed is not taken seriously, neither by students or by teachers.
2. Employers do value these achievements and often want a profile of applicants in terms of these achievements.
3. Governments often link these achievements with the knowledge economy, innovation and national well-being. They expect evidence that the achievements are promoted and assessed.
4. Judgements of these achievements help students and teachers identify ways of developing them further.

So, how might wicked competences, complex achievements or graduate attributes be assessed?

Tests of 'wicked' competences

Tests are widely used, especially by psychometricians, to 'measure' wicked competence. Let's explore what they do and don't do by looking at one, the TalentSmart emotional intelligence test (this and all following <http://www.eiquickbook.com/me/survey1of5.php>, password protected). It

... provides you with a complete picture of your emotional intelligence. This includes an understanding of:

- What emotional intelligence is
- Your overall emotional intelligence score
- Your current skill levels in the four areas that make up emotional intelligence
- Specific recommendations for action you can take to improve your emotional intelligence
-

It relies on self-reported typical behaviours. Questions ask 'how often (on a six-point scale from 'always' to 'never') you:

1. Are confident in your abilities.
2. Admit your shortcomings.
3. Understand your emotions as they happen.
4. Recognize the impact your behavior has upon others.
5. Realize when others influence your emotional state.
6. Play a part in creating the difficult circumstances you encounter.
7. Can be counted on.
8. Handle stress well.
9. Embrace change early on.
10. Tolerate frustration without getting upset.
11. Consider many options before making a decision.
12. Strive to make the most out of situations whether good or bad.
13. Resist the desire to act or speak when it will not help the situation.
14. Do things you regret when upset.
15. Brush people off when something is bothering you.
16. Are open to feedback.
17. Recognize other people's feelings.
18. Accurately pick up on the mood in the room.
19. Hear what the other person is 'really' saying.
20. Are withdrawn in social situations.
21. Directly address people in difficult situations.
22. Get along well with others.
23. Communicate clearly and effectively.
24. Show others you care what they are going through.
25. Handle conflict effectively.
26. Use sensitivity to another person's feelings to manage interactions effectively.
27. Learn about others in order to get along better with them.
28. Explain yourself to others.

Let's be clear that there is no reason to doubt that these items do, by normal psychometric standards, provide a robust description of self-reported behaviours, although test design is such that any sentient being would be able to 'fake good' by always choosing the 'always' response. Authors say that:

These 28 items derive from a view of emotional intelligence as: four important skills. The first two skills focus on you:

(1) Self-Awareness - Your ability to accurately perceive your emotions and stay aware of them as they happen. This includes keeping on top of how you tend to respond to specific situations and certain people.

(2) Self-Management - Your ability to use awareness of your emotions to stay flexible and positively direct your behavior. This means managing your emotional reactions to all situations and people.

The last two skills focus more on your contact with other people:

(3) Social Awareness - Your ability to accurately pick up on emotions in other people and get what is really going on. This often means understanding what other people are thinking and feeling, even if you don't feel the same way.

(4) Relationship Management - Your ability to use awareness of your emotions and the emotions of others to manage interactions successfully. Letting emotional awareness guide clear communication and effective handling of conflict.

The claim is that 28 questions are sufficient to determine that one of us (Knight) has the emotional intelligence of a teaspoon. His results were:

Your overall emotional intelligence score of 74 is higher than 46.0% of all people in the world.

Personal Competence

49.0% Self-Awareness Your score is higher than 49.0% of all people in the world.

48.0% Self-Management Your score is higher than 48.0% of all people in the world.

Social Competence

49.0% Social Awareness Your score is higher than 49.0% of all people in the world.

15.0% Relationship Management Your score is higher than 15.0% of all people in the world.

Some self-defence is that the extension of tests scores to comments about emotional intelligence is problematic and making inferences about behaviours is *deeply* problematic. This is notwithstanding the authors' credentials:

Dr. Travis Bradberry is a captivating speaker and coach, pushing the performance of people and organizations around the globe. His work has assisted leaders of Fortune 500 companies, all three branches of government including the US Senate, and even royalty abroad. He holds a dual doctorate in Clinical and Industrial-Organizational Psychology.

Dr. Jean Greaves is an award-winning consultant and entrepreneur with twenty years of experience helping organizations and people reach their potential. She specializes in challenging periods of rapid growth, with a third of the Fortune 500 having benefited directly from her expertise. She holds a doctorate in Industrial-Organizational Psychology.

Even so, the obvious question is whether self-report about behaviours in typical situation is a good basis for describing this 'wicked' competence. Were it to be shown that there was a good correlation between these scores and judgments made by other ways – if, for example, a latter-day Mahatma Gandhi scored perfectly while George W Bush scored 0 – then there would be reason to trust the scores. No such data about predictive validity exist. A more robust view of the value of this test would refer to the extensive literature on the many nuances of self-report.

Whatever the internal qualities of such tests – and it is quite credible that this test scores highly on Cronbach alpha and that little is added by adding more questions – it takes a step of faith to believe that highly-reliable test scores are a fair description of in-the-world inter-relationships.

Here, then, we have a good illustration of the difficulties with apparently reliable and valid tests of 'wicked' competences. There is no reason at all to doubt them, taken on their own terms.

Problems arise when we ask:

1. Is this test really a valid attempt to address the achievement or competence in question?
2. Does it accurately predict future performance or, at least, correlate highly with other ways of describing behaviour?

If the answer to either or both of these questions is 'no', then other means of assessing 'wicked' competences are needed.

2. The design of a study of the assessment of wicked competences

With advice from pbpl colleagues in three faculties of the Open University (OUBS, FELS and HSC) key informants were identified in six subject areas:

Accountancy
Early years teaching
Nursing
Secondary school teaching
Social work
Youth work

Key informants were asked to nominate up to five other informants working in other universities. All of these informants were contacted and asked, after being oriented to the terminology, to identify two 'wicked' competences that cause them assessment difficulties, be they cost, logistics, reliability, equal opportunities, or whatever.

Responses for each profession were labelled and aggregated. The labelling was simply a way of attaching a 'tag' to responses. These tags were created by the project officer and reviewed by the project director. In one case two 'tags' were collapsed under one heading ('Confidence – self reliance' and 'Self Management – awareness of personal effectiveness' became 'Self Management – confidence and effectiveness'). The temptation to bring other tags together into bigger conceptual groups was otherwise resisted.

Two competences from each area were targeted for close study. They were selected by (a) identifying those most commonly named by subject area informants (b) in case of ties, selecting competences that would extend the range of competences to be investigated. Nine were chosen for further study. Six were specific to one area and three were shared by two areas.

A web search was done to expand understanding of the current usage of each target competence in the associated subject areas.

Informants, as well as others recruited through Higher Education Academy subject networks and professional bodies, were asked, by on-line survey⁵, about assessment practices, purposes, problems and possibilities in respect of the target competences. There were 83 responses as at 20 November 2006. These data were quantitatively analysed.

A sub-sample of 14 was contacted by phone in December 2006 and January 2007 to elaborate points revealed by the analysis of the survey data. Interview data were thematically analysed following established qualitative methods, aided by Nvivo7 software.

⁵ The survey was available in three formats, including hard copy, but only on-line responses were received.

3. Nine wicked competences

The nine

Competence	Subject areas
a. Developing supportive relationships	Secondary teaching Youth work
b. Emotional intelligence	Youth work
c. Group work	Accounting
d. Listening and assimilating	Nursing
e. Oral communication	Accounting Early years teaching
f. Professional subject knowledge	Social work
g. Relating to clients	Nursing Secondary teaching
h. Self-management (confidence and effectiveness)	Early years teaching
i. 'Taking it onwards' – acting on diagnoses	Social work

This set of nine has some interesting features:

- They are heterogeneous. For example, emotional intelligence is presented as a form of intelligence or personal attribute. Professional subject knowledge is either declarative of propositional knowledge. The other seven could be called 'skills' and located with procedural knowledge ('know how') but finer-grained distinctions could be important: self-management falls close to Dweck's (1999) concept of self-theories; oral communication can be decomposed into various forms of speaking; and it is not all clear what developing supportive relationships means.
- Most are not expressed in the language of official standards, which tend to specify what professionals should be able to do. Listed here are competences that lie behind areas of performance. In secondary school teaching, the quality of classroom practice is associated with the quality of relationships; the quality of community nursing is associated with the quality of listening and hearing, and of relationships.
- Eight of them, at least, are complex in that it is not possible to specify in general, precise and unambiguous terms what any level of competence would be. In principle it might be possible to be more certain about professional subject knowledge.

Will clear criteria deal with assessment problems?

In exploring the assessment of 'wicked' competences, it is helpful to know that in the eyes of QAA there is only really a problem of will – of having the will to develop sufficiently clear criteria. The complexities and ambiguities to which we have referred in this report are seen by the QAA as ultimately resolvable, which is an interesting denial of Wittgenstein's point that rules cannot specify their own interpretation.

The Quality Assurance Agency's 2006 paper, *Outcomes from Institutional Audit: Assessment of students* said:

In the audit reports considered, the major themes emerging from the recommendations relating to assessment of students were:

- The development and use of consistent assessment policies
- The classification of undergraduate degrees
- The operation of assessment boards
- The security and reliability of assessments
- The use of assessment criteria
- The provision of feedback on students' work. (p 6)

Only in the section 'The use of assessment criteria' was there is any reference to issues surrounding the assessment of 'wicked' competences:

39 There are specific contexts in which the identification of assessment criteria has presented institutions with particular challenges. In one report (on an institution which faced difficulties in 'articulating standards in relation to such key attributes as "creativity" and "imagination"'), auditors learnt that the institution encouraged academic staff to adapt perceived good practice to their own contexts. They also learnt, however, that there was 'variable clarity of statements on assessment criteria' in course handbooks, and 'difficulties encountered by a final examination board in determining on what grounds a student's performance was to be judged unsatisfactory'. The institution was advised 'to expedite the formulation and implementation of institution-wide performance descriptors linked to assessment criteria which will provide a secure basis for the determination of examination outcomes'. (pp. 11 & 12)

The section refers to "Precept 7 of *Section 6* of the *Code of Practice*, which advises institutions to 'publish, and implement consistently, clear criteria for the marking and grading of assessments'. (p. 10). The implication is that doing so is relatively unproblematic – certainly there is no reference to significant problems in devising, communicating and applying criteria. The report also emphasises the value of a programme-wide approach to assessment:

Programme specifications typically set out intended learning outcomes which students are expected to achieve. Reports which draw attention to assessment criteria and methods have indicated that, for the benefit of students and examiners, it is helpful if they are explicitly aligned with these learning outcomes. (p.11)

Insofar as the QAA's audits see that there might be issues attaching to the assessment of wicked competences, they present a straightforward response – develop and apply clear criteria. Much the same advice is found in the work of American scholars such as Banta and colleagues (1996), Walvoord and Anderson (1998) and Maki (2004). QAA appears to attribute problems associated with the assessment of performance to deficient or absent assessment criteria. That can easily become a criticism of academic staff, who get blamed for failing to resolve what can be regarded as intractable problems with framing rules and the terms by which they are to be interpreted. As we look at the six subjects and nine wicked competence it is worth keeping this interpretation in mind and wondering whether it is possible to envisage clear criteria that allow robust generalisations to be made about 'wicked' competences, or whether the task is inherently not possible. In the former case there would need to be investment in the development of criteria and in training in their application. In the second, we would need to resign ourselves to judgements of 'wicked' and suchlike competences having no more than local meanings.

4. The six subject areas: wicked competences and their assessment

This section comprises notes on the two wicked competences chosen for each subject area and on their assessment in that subject area. Where possible, the notes contain extracts from the QAA subject benchmark statements, produced at the turn of the century. Failing that, national standards are used, as in the case of early years and secondary phase teaching. In the case of Youth work, it was necessary to take data from a national subject association, since the National Training Organisation responsible for the field was disbanded along with other NTOs and there not yet appear to be a replacement for it.

Accounting

Responses from informants identified two 'wicked' competences for attention. *Group work* and *oral communication*. The UK Quality Assurance Agency benchmark statement for 'accounting' (QAA, 2000a, <http://www.qaa.ac.uk/academicinfrastructure/benchmark/honours/accounting.asp>, accessed 17 August 2006) have several relevant points to make about the assessment of these competences (or 'skills' as the QAA calls them).

2.2. Accounting as a degree subject requires students to study how the design, operation and validation of accounting systems affects, and is affected by, individuals, organisations, markets and society. This study is informed by perspectives from the social sciences. Such perspectives may include, but are not restricted to, the behavioural, the economic, the political, and the sociological. As indicated above, accounting is often studied in combination with a substantial amount of finance. Under such circumstances, the degree structure should also require the study of the operation and design of financial systems, risk, financial structures, and financial instruments. (p. 1)

4.1. On completion of a degree programme covered by this statement, a student should have acquired the following abilities and skills:

- vii. communication skills including the ability to present quantitative and qualitative information, together with analysis, argument and commentary, in a form appropriate to the intended audience;
- viii. normally, ability to work in groups, and other inter-personal skills, including oral as well as written presentation skills. (p. 2)

5.4. No single form of assessment activity is uniquely appropriate for evaluating student achievement on degree programmes in accounting. There should be a suitable balance and mix of assessment activities to allow and require students to demonstrate not only their understanding of the conceptual and applied aspects of accounting but also the cognitive abilities and non-subject-specific skills they have developed as a consequence of their studies. Also, consideration needs to be given to the balance between formal assessment activities and other forms of non-assessed experiences that together contribute to the development of an accounting graduate. (p. 3)

6. 5. Threshold graduates ... will demonstrate possession of the required cognitive abilities and non-subject specific skills to a basic level of achievement. [*Cognitive abilities and skills* - basic levels of attainment are characterised by minimal proficiency in the ability or skill. Graduates with a threshold level of attainment can be expected perform well in simple or straightforward situations.] (pp. 4, 3)

6.7. Typical graduates can distinguish themselves from threshold graduates by displaying a more thorough knowledge and understanding and enhanced technical abilities. They can also demonstrate an enhanced capacity to develop and apply critical, analytical and problem solving abilities and skills. (p. 4)

Early years education

Responses from informants identified two 'wicked' competences for attention: *self management* (confidence and effectiveness) and *oral communication*. The UK Training and Development Agency for Schools' *Professional Standards for Qualified Teacher Status* (<http://www.tda.gov.uk/upload/resources/doc/q/qualifying-to-teach-in-microsoft-word-fo.doc>, accessed 17 August 2006) has relevant points to make about the assessment of these competences. People who meet the standards are ones who:

S1.7 They are able to improve their own teaching, by evaluating it, learning from the effective practice of others and from evidence. They are motivated and able to take increasing responsibility for their own professional development. (p. 8)

S3.2.7 They are able to use records as a basis for reporting on pupils' attainment and progress orally and in writing, concisely, informatively and accurately for parents, carers, other professionals and pupils. (p. 12)

S3.3.11 They can take responsibility for teaching a class or classes over a sustained and substantial period of time. They are able to teach across the age and ability range for which they are trained. (p. 14)

This is a very good example of a set of standards that concentrates on 'front-of-house' performances, leaving underpinning competences to be inferred. Plainly an early years teacher must have oral facility and be very skilled in all forms of management. They are simply taken for granted in standards that refer to the performances to which such competences contribute.

The Qualifications and Curriculum Agency, which designs and monitors the national curriculum, has very little to say about teachers, beyond this in its 2004 report on the Foundation Stage (ages 3-5).

Factors affecting the quality of provision

5. Recruitment and retention of staff, particularly for settings in the non-maintained sector, can adversely affect the quality of the curriculum offered. Poor pay, long working hours and significant increases in accountability and paperwork were the main reasons cited for these recruitment and retention difficulties.

6. There is significant variation in staffing ratios across both the non-maintained and maintained sectors. Practitioners consider that inadequate adult to child ratios have a significant impact on the quality of the provision they are able to offer. They make it difficult to provide children with access to indoor and outdoor learning simultaneously. This is particularly so for reception classes. Practitioners called for a nationally agreed minimum adult to child ratio across the foundation stage. (p. 5)

Ofsted, which inspects the quality of teaching and is often critical, has said nothing in recent years about general teaching competences, concentrating on passing general verdicts about

teaching (satisfactory, not satisfactory) or on teaching specific subjects (especially reading, science and maths).

Nursing

Responses from informants identified two 'wicked' competences for attention: *listening & assimilating* and *relating to clients*. The Nursing and Midwifery Council's undated Standards of proficiency for Pre-registration Nursing Education (<http://www.nmc-uk.org/aFrameDisplay.aspx?DocumentID=328>, accessed 17 August 2006) have quite a lot to say about relating to clients. Nurses should:

- Practise in a fair and anti-discriminatory way, acknowledging the differences in beliefs and cultural practices of individuals or groups.
- Engage in, develop and disengage from therapeutic relationships through the use of appropriate communication and interpersonal skills.
- Create and utilise opportunities to promote the health and well-being of patients, clients and groups.
- Undertake and document a comprehensive, systematic and accurate nursing assessment of the physical, psychological, social and spiritual needs of patients, clients and communities.
- Formulate and document a plan of nursing care, where possible in partnership with patients, clients, their carers and family and friends, within a framework of informed consent. (p. 5)

Respect for individuals and communities

All members of the profession must demonstrate an inviolable respect for persons and communities, without prejudice, and irrespective of orientation and personal, group, political, cultural, ethnic or religious characteristics. Care must be provided without prejudice and in an antidiscriminatory fashion. No member of the profession should convey any allegiance to any individual or group affiliations which oppose or threaten the human rights, safety or dignity of individuals or communities, irrespective of whether such individuals or groups are recipients of care. (p. 16)

Students should:

- Discuss methods of, barriers to, and the boundaries of, effective communication and interpersonal relationships
- Recognise the effect of one's own values on interactions with patients and clients and their carers, families and friends
- Utilise appropriate communication skills with patients and clients
- Acknowledge the boundaries of a professional caring relationship.
- Demonstrate sensitivity when interacting with and providing information to patients and clients. (27)

Only by implication is 'listening and assimilation' present, mainly in the reference to undertaking 'a comprehensive, systematic and accurate nursing assessment of the physical, psychological, social and spiritual needs of patients, clients and communities'. The 'key skills' section has nothing to say about listening, nor does any other section of the standards.. The closest is:

Demonstrate literacy, numeracy and computer skills needed to record, enter, store, retrieve and organise data essential for care delivery (p. 33, emphasis in original)

Secondary school education

Responses from informants identified two 'wicked' competences for attention: *developing supportive relationships* and *relating to clients*. In the event, respondents chose not to comment on the second of these secondary school education competences. The UK Training and Development Agency for Schools' *Professional standards for qualified teacher status* (<http://www.tda.gov.uk/upload/resources/doc/q/qualifying-to-teach-in-microsoft-word-fo.doc>, accessed 17 August 2006) have several relevant points to make about the assessment of these competences. People who meet the standards are ones who:

S1.1 They have high expectations of all pupils; respect their social, cultural, linguistic, religious and ethnic backgrounds; and are committed to raising their educational achievement.

S1.2 They treat pupils consistently, with respect and consideration, and are concerned for their development as learners.

S1.3 They demonstrate and promote the positive values, attitudes and behaviour that they expect from their pupils.

S1.4 They can communicate sensitively and effectively with parents and carers, recognising their roles in pupils' learning, and their rights, responsibilities and interests in this. (p. 8)

S2.4 They understand how pupils' learning can be affected by their physical, intellectual, linguistic, social, cultural and emotional development. (p. 10)

S3.1.2 ... They take account of and support pupils' varying needs so that girls and boys, from all ethnic groups, can make good progress.

S3.1.3 They select and prepare resources, and plan for their safe and effective organisation, taking account of pupils' interests and their language and cultural backgrounds, with the help of support staff where appropriate. (p. 11)

S3.3.4 They differentiate their teaching to meet the needs of pupils, including the more able and those with special educational needs.

S3.3.6 They take account of the varying interests, experiences and achievements of boys and girls, and pupils from different cultural and ethnic groups, to help pupils make good progress.

S3.3.9 They set high expectations for pupils' behaviour and establish a clear framework for classroom discipline to anticipate and manage pupils' behaviour constructively, and promote self-control and independence. (p. 14)

The General Teaching Council's *Statement of Professional Values and Practice for Teachers* (2006, <http://www.gtce.org.uk/standards/disc/StatementOfProfValues>, accessed 21 August 2006) also refers to the quality of relationships in teaching:

Children and young people

Teachers treat young people fairly and with respect, take their knowledge, views, opinions and feelings seriously, and value diversity and individuality. They model the characteristics they are trying to inspire in young people, including enthusiasm for learning, a spirit of intellectual enquiry, honesty, tolerance, social responsibility, patience, and a genuine concern for other people.

Parents and carers

Teachers respond sensitively to the differences in the home backgrounds and circumstances of young people, recognising the key role that parents and carers play in children's education.

They seek to work in partnership with parents and carers, respecting their views and promoting understanding and co-operation to support the young person's learning and well-being in and out of school.

Professional colleagues

Teachers see themselves as part of a team, in which fellow teachers, other professional colleagues and governors are partners in securing the learning and well-being of young people.

They recognise the importance of effective multi-agency working, are clear and confident about their own role and professional standards, and understand and respect the roles and standards of other colleagues. They are keen to learn from others' effective practice and always ready to share their own knowledge and expertise. They respect young people's and colleagues' confidentiality wherever appropriate.

It is clear from both sets of statements that secondary teachers must be good at relationships but it is implicit in descriptions of the performances to which such competences contribute.

Social work

Responses from informants identified two 'wicked' competences for attention, *professional subject knowledge* and *'taking it onwards'*. The former is relatively well-defined, albeit at some length. The latter is understood here in terms of acting out the analysis and decision-making aspects of problem-solving.

The UK Quality Assurance Agency benchmark statement for social work (QAA, 2000b, <http://www.qaa.ac.uk/academicinfrastructure/benchmark/honours/socialpolicy.asp>, accessed 17 August 2006) have several relevant points to make about the assessment of these competences (or 'skills' as the QAA calls them).

2.1 As an applied academic subject, social work is characterised by a distinctive focus on practice in complex social situations to promote and protect individual and collective well-being. At honours degree level the study of social work involves the integrated study of subject specific knowledge, skills and values and the critical application of research knowledge from the social and human sciences (and closely related domains) to inform understanding and to underpin action, reflection and evaluation. Honours degree programmes should be designed to help foster this integration of contextual, analytic, explanatory and practical understanding. The specific areas of knowledge and understanding and the relevant subject and other skills to be acquired are defined in sections 3.1 and 3.2 of this statement. (p. 11)

2.2.4 It follows that, through their education, honours graduates in social work should be:

- equipped both to understand, and to work within, this context of contested debate about nature, scope and purpose;
- enabled to analyse, adapt to, manage and eventually to lead the processes of change. (p. 11)

2.3 The applied nature of social work as an academic subject means that practice is an essential and core element of learning. (p. 11)

3.1 Subject knowledge and understanding

- During their degree studies in social work, honours graduates should acquire, critically evaluate, apply and integrate knowledge and understanding in the following five core areas of study:
 - 3.1.1 Social work services and service users [5 paragraphs follow]
 - 3.1.2 The service delivery context [8 paragraphs follow]
 - 3.1.3 Values and ethics [5 paragraphs follow]
 - 3.1.4 Social work theory [6 paragraphs follow]Approaches and methods of intervention in a range of community-based settings including group-care at individual, group and community levels, including factors guiding the choice and evaluation of these.
 - Knowledge and critical appraisal of relevant social research and evaluation methodologies.- 3.1.5 The nature of social work practice [6 paragraphs follow] (pp 12-14)
- 3.2.2 Problem solving skills
These are sub-divided into four areas:
 - 3.2.2.1 *Managing problem-solving activities*
Honours graduates in social work should be able to plan problem-solving activities, i.e. to:
 - think logically and systematically;
 - apply ethical principles and practices critically in planning problem-solving activities;
 - plan a sequence of actions to achieve specified objectives;
 - manage the processes of change ...
 - 3.2.2.4 *Intervention and Evaluation*
Honours graduates in social work should be able to use their knowledge of a range of interventions and evaluation processes selectively to:
 - build and sustain purposeful relationships with people and organisations in community-based, and inter-professional contexts including group-care;
 - make decisions, set goals and construct specific plans to achieve these, taking into account relevant factors including ethical guidelines;
 - negotiate goals and plans with others, analysing and addressing in a creative manner human, organisational and structural impediments to change;
 - implement plans through a variety of systematic processes including contracting with others;
 - undertake practice in a manner that promotes the well-being and protects the safety of all parties;
 - manage the complex dynamics of dependency and, in some settings, provide direct care and personal support in every day living situations;
 - meet deadlines of time and comply with external definitions of task;
 - monitor situations, review processes and evaluate outcomes;
 - bring work to an effective conclusion, taking into account the implications for all involved. (pp. 14, 15)
- 4.4 Assessment strategies should be chosen to enhance students' ability to conceptualise, compare and analyse issues from a range of data sources including practice and their capability to practise. Academic assessment is designed to develop and test cognitive skills, drawing on the contexts of practice and reflecting the learning and teaching methods employed. Methods normally include case study presentations and analyses, practice-focused assignments, essays, project reports and examinations. The requirements of honours degree programmes in social work frequently include an extended piece of written work, which may be practice-based, and is typically undertaken in the final year. This may involve independent study for either a dissertation or a project, based on systematic enquiry and investigation.

4.5 Honours degree programmes in social work assess practice not as a series of discrete practical tasks, but as an integration of skills and knowledge with relevant conceptual understanding. This assessment should, therefore, contain elements that test students' reflective analysis. Where the honours degree is an integrated academic and professional award, the failure of any core element including assessed practice will mean failure of the course. (p. 17)

Standards associated with three distinct levels of attainment are identified below. Given the essentially applied nature of social work, standards are specified in relation to *both* academic and practice capabilities. At each level, the requirements relate to subject specific knowledge, understanding and skills (including key skills inherent in the concept of 'graduateness'). Students will be expected to meet each of these requirements. Where there is an integration of honours degree and professional requirements, meeting these standards should enable students broadly to meet the requirements of the regulatory bodies.

5.1 Modal level

This level represents that of the typical student graduating with an honours degree in social work. At modal level, students will be able to demonstrate the following:

5.1.1 Knowledge and understanding

- Sound understanding of the five core areas of knowledge and understanding relevant to social work as detailed in section 3.1, including their application to practice and service delivery;
- ability to use this knowledge and understanding in an integrated way in specific practice contexts;
- appraisal of previous learning and experience and ability to incorporate this into their future learning and practice;
- acknowledgement and understanding of the potential and limitations of social work as a practice based discipline;
- ability to use research and enquiry techniques with reflective awareness to collect, analyse and interpret relevant information; and
- developed capacity for the critical evaluation of knowledge and evidence from a range of sources.

5.1.2 Subject skills and other skills

A developed capacity to:

- apply creatively a repertoire of core skills as detailed in section 3.2;
- integrate clear understanding of ethical issues and codes of values and practice with their interventions in specific situations;
- consistently exercise an appropriate level of autonomy and initiative in individual decision making within the context of supervisory, collaborative, ethical and organisational requirements; and
- demonstrate habits of critical reflection on their performance and take responsibility for modifying action in light of this. (pp 17, 18)

Youth work

Responses from informants identified two 'wicked' competences for attention: developing supportive relationships and emotional intelligence. There are no UK Quality Assurance Agency benchmark statements for Youth Work and the national Training Organisation, PAULO, is being replaced and its web site was broken on 17 August. *The NYA Guide to Youth Work and Youth Services* was used as an alternative point of reference about the position of these competences in youth work practice.

One complication is the range of titles and settings that attach to 'youth work':

Since youth and community workers work in a wide range of settings, higher education qualifications reflect different occupational needs, and have a range of titles, including youth and community, community and youth studies, childhood and youth studies, and informal and community education. (p. 27)

The competences youth workers need have to be inferred from sections of the NYA booklet including:

Youth workers are not expected to be an expert on everything affecting young people. But they do need to be aware of other local agencies and what they can offer young people, and to recognise when they need to involve people with specialist skills and knowledge, while continuing to support the young person concerned. As young people come to trust workers, they may confide details of difficult personal circumstances. In such situations it is important that the youth workers and young people are both clear about when such information will be kept confidential, and when it may have to be passed on to other people. (p. 22)

"A session for young people leaving care at the 16 plus Team in Swanley elicited deeply held views and the worker was able to console, inspire, guide and inform. Youth workers often used humour to challenge young people's low horizons and raise their self-esteem. They allowed space for, and welcomed the involvement of, students in preparing sessions as well as in the practical tasks of setting up equipment and taking responsibility for outings." Ofsted, Kent, 2002. (p. 25, emphasis in original)

"Centrepont Youth Club was an outstanding example of provision for young people with disabilities where youth work was centred on the belief that all young people could achieve their aspirations and targets. The empowerment of individuals was at the core of this work and the success so significant that the members themselves felt confident enough to take over the management and leadership of the club shortly, and the workers were brave enough to feel they could let go." Ofsted, Somerset 2002. (p. 26, emphasis in original)

5. The competences

To some degree this section re-visits the ground covered in the last section, in which subject-specific references to these wicked competences and their assessment were set out. A problem with that treatment is that words are used, not defined, and we may be scarcely more certain about their meaning after reading through professional standards documents than we were beforehand. This section reports web searches designed to help us understand what these competences are – what the words describing them *mean*.

a. Developing supportive relationships

Intuitively, we know that this is an important competence in many professions. Surprisingly, a web search 'skill "developing supportive relationships"' had only 44 hits. None described the skill or attributed any meaning to the term – the target documents simply used the phrase in a variety of connections.

The nursing and secondary phase standards emphasise the importance of developing a variety of supportive relationships but do not provide much help in understanding what it involves, nor how it might be recognised and assessed.

b. Emotional intelligence (EI)

An on-line search found about a quarter of a million 'hits' for 'skill "emotional intelligence"'. Wikipedia (http://en.wikipedia.org/wiki/Emotional_intelligence#Ability-based_measures_of_EI, accessed 18 August, 2006) says:

Current definitions of EQ are inconsistent about what it measures: some (such as Bradberry and Greaves 2005) say that EQ is dynamic, it can be learned or increased; whereas others (such as Mayer) say that EQ is stable, and cannot be increased ... The Mayer-Salovey (1993) model defines emotional intelligence as the capacity to understand emotional information and to reason with emotions. More specifically, they divide emotional intelligence abilities into four areas -- in their four branch model:

1. The capacity to accurately perceive emotions.
2. The capacity to use emotions to facilitate thinking.
3. The capacity to understand emotional meanings.
4. The capacity to manage emotions.

These four abilities are assessed by criterion-based (or abilities-based) tests (the researchers have introduced several versions, the latest of which is the MSCEIT V2.0)

There is a view that the construct of EI may not stand scrutiny.

- Some say that it may actually be measuring conformity;
- Eysenck (2000), for example comments that Goleman "exemplifies more clearly than most the fundamental absurdity of the tendency to class almost any type of behaviour as an 'intelligence'. . . . If these five 'abilities' define 'emotional intelligence', we would expect some evidence that they are highly correlated; Goleman admits that they might be quite uncorrelated, and in any case if we cannot measure them, how do we know they are related? So the whole theory is built on quicksand; there is no sound scientific basis."
- There is a view that self-report EI may merely be another measure of personality.

If the concept itself is not robust, then it is hard to see how it is to be assessed. Nevertheless, scientists have produced several 'measures', one of which was reviewed earlier in this chapter. A major test of one interpretation of EI is MSCEIT:

What does emotional intelligence, and the MSCEIT, predict? Contrary to the claims in the popular press, we are certain that emotional intelligence is not "twice as important as IQ". Indeed, we know of no psychological variable that is that powerful a predictor. The

MSCEIT will likely predict important outcomes, but at levels that one usually obtains in psychological research.

There are a number of studies that are in the field, but those that have been completed suggest that the MSCEIT offers additional predictive validity for outcomes such as pro-social behavior, deviancy, and academic performance. (Mayer, Salovey, Caruso Emotional Intelligence Test, 2005, <http://www.emotionaliq.org/MSCEIT.htm>, accessed 18 August 2006)

Another test is the EQi, created by Reuven BarOn and published by Multi-Health Systems, also the publisher of the MSCEIT. Other assessments include the EQ Map, the Six Seconds Emotional Intelligence Assessment, the Emotional Competence Inventory, and the Ei360.

Tett, Fox, and Wang (2005) provide a psychometric review of 6 self-report measures of emotional intelligence based on 33 peer-reviewed reports. More importantly, it introduces a psychometrically rigorous measure of trait-based emotional intelligence based on (Salovey and Mayer 1990). (Wikipedia, http://en.wikipedia.org/wiki/Emotional_intelligence#Ability-based_measures_of_EI, accessed 18 August 2006)

The difficulty is that the assessment devices are addressing different aspects of the construct and judging them in different ways. Consequently, there is no clear steer on how this 'wicked' competence might be assessed.

c. Group work

A web search for 'skill "group work"' returned nearly half-a-million pages. In those visited, the majority were references to work being done in groups or declarations of the importance of being able to work in groups. As such, they were uninformative. There were some descriptions of what you need to do to work well in groups. A typical example comes from Liverpool Hope University's Geography Department (<http://www.hope.ac.uk/gnu/stuhelp/groupwk7.htm#Ten%20'Do%20s'%20and%20'Don'ts'>, accessed 17 August, 2006).

Ten 'Dos' and 'Don'ts' for a student working in a group. (compiled using comments made by current first year Geography undergraduates at Liverpool Hope)

DO:

1. Contribute to the group by putting forward ideas and getting involved in discussion.
2. Co-operate with other members of the group.
3. Select specific roles for every group member (e.g. Spokesperson, Leader, Time-keeper, Scribe etc.)
4. Take on your given role within the group.
5. Organise meeting times and stick to them.
6. Share the workload equally.
7. Respect the values and opinions of others.
8. Communicate and share ideas with each other.
9. Encourage others to speak.
10. Value diversity.

DON'T:

1. Sit back and do nothing.
2. Think that others will do the work.
3. Take lead and ignore others.
4. Be afraid to speak.
5. Expect all your own ideas to be used.
6. Leave people out of the discussion.
7. Become distracted from the subject.

8. Be aggressive, interrupt or criticise unnecessarily.
9. Ignore other members of the group.
10. Allow one person to dominate.

A presentation on assessing group work

(http://www.brad.ac.uk/acad/tqeg/resources/assessing_group_work/live_assess_grp_work.ppt#271,1, Assessing Group Work, accessed 20 August, 2006) is in fact how to assess the quality of group work achievements, not how to assess individuals' performances when working in groups. Fair assessment of group productions is a recurrent theme in the literature, especially because it is common to derive individual marks from group activities. The following extract from the University of Durham is typical of guidance on fair practice:

Group Work

The obvious problem in assessing group work is that of being fair to the individuals within the group, not all of whom may have made an equal contribution to the group task. If group work is to be assessed summatively the following examples of good practice should be considered:

1. ensure that there is a clear, subject-related 'product' or 'outcome' for the group work which can be assessed as 'content' in addition to assessment of the group 'process' (how well the group worked together).
2. have a tight marking scheme and clear criteria for the assessment of the 'content' of the group work ...
3. ensure that the work of each individual can be identified within or alongside the group work, e.g.:
 - i. an individual summary of the work from each student;
 - ii. identification by the group of the contribution of each person;
 - iii. the group does the 'ground-work' but each individual produces his/her own work from that;
 - iv. a research diary or field notebook kept by each individual;
 - v. a piece of work from each individual associated with the group work to show that s/he has a grasp of the issues involved;
 - vi. a piece of work from each individual evaluating the group process as well as the output;
4. have some sort of moderation of the mark for the piece of work to take into account the variable contribution of individuals. This can be done by:
 - i. allocating the same mark to all group members but allowing students to 'appeal' against it where they do not feel that it accurately reflects the contribution of each individual;
 - ii. allowing the students to allocate a percentage of the marks within the group for the contribution of each person;
 - iii. deducting marks for clear-cut examples of non-contribution such as failure to attend group meetings;
 - iv. obtaining feedback from students on how the individual group members performed which the tutor uses to reallocate marks to reflect the contribution of each member;
 - v. using the tutor or supervisor as a 'consultant' to deal with problems faced by the group so that s/he is aware of the levels of contribution of each group member.

(<http://www.dur.ac.uk/teachingandlearning/handbook/6-1-4-16.pdf#search=assess%20skill%20group%20work>, accessed 18 August 2006)

There are obviously problems in trying to assess without descriptions of the nature of effective group work. Tempting though it is to use 'dos and don'ts' as the basis for judgement, there is no evidence that these are sound indicators of effective group work, let alone that they are invariant. There are more authoritative sources, such as Jacques (2000) or Belbin (2003). Belbin, for example, identifies eight types of teams and nine role types. There is an on-line service

administering the Belbin self-perception inventory. For £25.00 per person 'you will receive (among other reports) a 'fingerprint' of your preferences for each of the 9 Team Roles' (<http://www.belbin.com/testing.htm>, accessed 21 August, 2006). There seem to be several assessment problems here, the most challenging of which is the complexity of the concept – with such a high degree of variety in the concept, assessment would involve many observations in varied teams, roles and settings. While it may be that Belbin's assessment service provides robust descriptions of group work achievement, it is hard to see that this would be a reasonable solution for university courses, if only on grounds of cost.

d. Listening and assimilation

The search 'skill "listening"' had over six million hits. ('Assimilating was dropped from the search because the phrase 'listening and assimilating' was found to be too esoteric.) A large proportion of the hits seemed to relate to foreign language learning but there remained a good number potentially relevant to listening as a professional practice. Closer inspection found many examples focusing on young children and others with specific learning needs. Not untypical of that set is:

SKILL: Listening - Attend to speaker and acknowledge what is said.

STEPS :

1. Face person who is talking-"eye contact."
2. Sit quietly.
3. Attend to what is being said.
4. Wait your turn to talk.

(<http://www.users.globalnet.co.uk/~ebdstudy/strategy/socskills.htm>, accessed 21 August 2006)

ManageTrainLearn (<http://www.managetrainlearn.com/live/al8.asp>, accessed 21 August 2006) reckons that focused, high-concentration, razor-sharp listening is the most important skill for any manager to possess. Apart from some advice on the mechanics of listening:

1. don't do any important listening if you've not got the time
2. remove any interference or localized noise
3. try to position yourself in front of the other person so that you can hear them clearly and see their lips
7. don't do any listening if your mind's on something else

There are sections on:

2. The Right Way To Listen and The Wrong Way
3. Empty The Cup of Your Own Prejudices
4. Master The Mechanics of Maximum Listening

There are 3 tricks that can help you stay on track and they're known as the ABC of maximum listening. They are...

A for Attending.

B for Body Language.

C for Conversation Encouragers.

5. Tune In To What's Really Being Said, Not Just What You Hear
6. Focus On Them, Not You
7. Hit the Bullseye With Empathic Listening

So those are the 7 keys to great listening. Listening that will make you a great communicator with the power to transform people's lives. Not transform, as in force to change, but transform as in "assist with what needs to happen."

An approach such as that suggested by ManageTeachLearn would cover 'assimilating' as well as 'listening'. As for assessment, problems might be anticipated because of the extensiveness of the concept as ManageTeachLearn presents it.

It was established that assessment is a major issue amongst those concerned with 'listening'. For example, The International Listening Association's 2006 conference, *Measuring the Effects of Listening on Learning, Earning, and Healing*, said that

... Participants will actively identify the challenges of measuring listening and design pilot studies to overcome the challenges. Pilot studies will be conducted by conference participants. The current editor of the International Journal of Listening has agreed to focus a special edition of the journal that will include conference outcomes and the pilot studies. Further publication opportunities will be available to conference participants as well.

Eight standardised listening tests are described at http://www.listen.org/Templates/listening_tests_instructor_manual.htm and there is a discussion of three main types of listening test by Rhodes et al., (1990) (<http://www.cios.org/www/ijl/ijl04toc.htm>, accessed 21 August 2006). Apart from raising fundamental issues about reliability, validity and good technical practice, which will be considered further in section 6 of this chapter, they make some useful general points. They identify problems with conceptualising listening. On the one hand there has been some disagreement as to whether it is an internal process and therefore beyond observation, or whether it includes the listener's responses, in which case observation is possible. On the other, it is not easy to draw the boundaries between listening, reading, communication and conversation. This is especially important because the majority of tests they mention seem to be for school-aged children. There may be significant differences between a kindergarten child listening in school and a professional listening to a client. They observe that 'it is impossible to assess a construct that cannot adequately be defined' (p. 66) and show that tests do not necessarily meet technical standards of reliability and validity and that there would be operational difficulties in trying to use the better tests on any large scale. An interesting section on 'informal assessment' steps aside from the measurement paradigm of the summative tests. Nowadays it might be seen as a section on formative or low stakes assessment, a topic to which we return in section 8.

Bostrom and colleagues' (2004) review of a listening comprehension test within the US Law School Aptitude Test (LSAT) is a serious piece of work informed by first-rate psychology, which is one reason why it draws attention to the role of working memory in listening. They also distinguish between 'lecture' or 'passive' listening and 'dialogic' listening, where there is a need to understand what was said and formulate a response (p. 11). An implication is that different tests may favour different sorts of listening. A strong finding is that although the LSAT listening comprehension test is viable,

.... We have questions about the construct being measured, and, as the developers themselves have shown, the results correlate too highly with both R[eadings] C[omprehension] and L[ogical] R[easoning] tests. Hence, given the cost of creating the LC items, which call for audio recordings of monologues and dialogues and for their presentation in the test itself, it would not seem sensible to include the current LC test in the LSAT. Comparable information about the candidates is already assessed in the RC and LR tests. (p. 4).

e. Oral communication

Web searches suggested that tests for 'oral communication' which don't involve learning languages are thin on the ground. There are courses, one of which is, unusually, to help university teachers in New York State to communicate their knowledge verbally (<http://www.hamilton.edu/academics/courses.html?dept=Oral%20Communication>). Hamilton University's site includes an article on 'how to take an oral exam', which has some useful hints and tips. However, there is no sign of an aptitude test to show how well a person is able to communicate orally. What seems to be emphasized at this site and in the other sites visited is that oral communication and listening skills are closely linked. Although this may seem obvious,

from an assessment point of view it is important to clear whether oral communication is conceived as clear transmission or as sensitive conversation.

The following extract from the University of Durham is typical of guidance on fair assessment practice:

Oral Communication Skills

1. Most of the assessment of oral communication skills practised in the University is integrated with tutorial or seminar sessions or with project work (a 'poster' with an oral defence, a viva or a presentation). In these cases the 'subject content' usually constitutes more than 50% of the mark for the presentation which enables the academic element of the assessment to predominate in the assessment.
2. It is essential that students know what proportion of their mark will depend on content and what on process and that the criteria for each aspect are made clear.
3. It is good practice to give students input about the process on which they are engaged - i.e. presentation skills. This may be delivered by an external specialist brought in by the department if colleagues do not feel confident about delivering this themselves. This should support the definition of the criteria on which the presentation skills will be assessed.
4. Moderation by the external examiner can be difficult or impossible, depending on the context in which oral skills are being assessed. Depending on the situation it may be possible for:
 - i. the presentation to be taped (this has been common practice for orals in modern languages for some time);
 - ii. final-year vivas administered by the external examiner to be used to assess, or moderate assessment of, communication skills;
 - iii. more than one member of staff to assess each presentation. It is good practice to have clear criteria and to use marking sheets consistently against those criteria so that the external can have sight of the marking notes;
 - iv. a third colleague to moderate all the presentations across an area in which s/he specialises;
 - v. students to provide peer assessment of presentations to complement staff assessment - this may also act as a catalyst for a discussion of 'what makes a good presentation and how to do it'.

(<http://www.dur.ac.uk/teachingandlearning.handbook/6-1-4-16.pdf#search='assess%20skill%20group%20work'>, accessed 18 August 2006).

f. Professional subject knowledge

Although this 'competence' is being associated in this study with the social work profession, it was also mentioned by nursing professionals. In both cases there is a great deal to know – regulations, procedures and evidence-based information about actions. In both subject areas there is a great deal of propositional knowledge to hold, which raises the technical issue of how assessments can be designed to sample the domain fairly without being so intrusive that they squeeze out other aspects of the professional programme.

The category of 'professional subject knowledge' is neither a clear nor a simple one. Shulman (1986), writing about schoolteaching, said that there is a need for content knowledge (which is a form of 'propositional knowledge'), for pedagogical knowledge (which is a form of procedural knowledge) and for pedagogical subject knowledge, which is a mix of propositional and procedural knowledge. Amongst the many assessment implications, one stands out: it will not be sufficient to assess propositional knowledge as a proxy for professional subject knowledge. Dreyfus and Dreyfus (2005) argue that there are marked differences in the ways in which experts

and novices organise and use their professional knowledge. The assessment implication is that different assessment approaches could be called for at different points on the novice-expert continuum. Swanwick (2005), writing of medicine, provides a useful summary of the diversity of knowledge that professionals need. Diverse knowledge implies diverse assessment practices.

The major assessment problem is that in these six subject areas professional knowledge is needed for practice in a way that is not the case in, for example, pure science, arts and social science subjects. The main assessment problem is whether to judge professional subject knowledge through performances⁶ (there are many practical problems attaching to this approach) or through proxies – paper or on-line tests of knowledge are proxies for assessments of knowledge-in-action. On-line and paper tests are efficient and convenient ways of assessing 'free-standing knowledge' but there is a view that professional knowledge should only be valued if it is properly deployed in practice settings. Deployment of knowledge is a form of transfer of learning and it is notorious that transfer does not happen easily or appropriately. People who score highly in on-line tests of knowledge may fail to transfer it to practical situations and therefore be, to all intents and purposes, ignorant.

g. Relating to clients

A web search 'skill "relating to clients"' had 410 hits. None of those visited described the skill or attributed any meaning to the term – the target documents simply used the phrase in a variety of connections.

h. Self-management

Wikipedia (<http://en.wikipedia.org/wiki/Self-management>, accessed 19 August 2006) discerns four applications of the term, the most relevant to our purposes being:

Self-management means different things in different fields:

- In business, education, and psychology, self-management refers to methods, skills, and strategies by which individuals can effectively direct their own activities toward the achievement of objectives, and includes goal setting, planning, scheduling, task tracking, self-evaluation, self-intervention, self-development, etc. Also known as *executive processes* (in the context of the *processes of execution*).

To some degree the encyclopaedia answers one question, namely whether the construct of 'self-management' has any validity, or whether it is just a convenient fiction. Wikipedia's is an extensive definition that creates assessment problems. As with 'listening' there is a definitional problem regarding the degree to which self-management is an internal process or an external one and, similarly, whether it is reasonable to assess self-management through self reports or through observations of behaviour. Taking any one of these four positions (internal, external, proxies, observation) implies quite different assessment problems and strategies.

i. 'Taking it onwards' – acting on diagnoses

Predictably, neither searching for 'skill "taking it forwards"' nor for 'skill "acting on diagnosis"' produced any hits. This competence is definitely about acting but where 'taking it forward' and 'acting on diagnoses' are too specific, 'acting' is hopelessly wide as a search term.

Here is the major assessment issue: how is the competence to be defined? And, in view of the absence of a general definition, is there any possibility of any definition or operationalisation having meaning outside the local context in which it was created?

⁶ There are many practical problems attaching to this approach – problems that were a major cause of the demise of the early 1990s approach to National Vocational Qualifications (Jessup, 1991)

6. Reprise: issues in the assessment of wicked competences

Generic issues that have been mentioned in these notes on the nine competences and six subject areas include:

1. Metaphysics

There is a tendency to act as if the competences in question are real, which implies that they can be measured. 'Measurement' is often used as a synonym for 'assessment' and 'measure' for 'assess'. There is an alternative ontology which draws attention to the fluxional and situated character of the human world. It follows that there is no method that allows us to be certain about it. In Law's (2004: 143) words

The argument has been that method in social science (and in natural science too) is enacted as a set of nineteenth- or even seventeenth-century Euro-American blinkers. This means that it misunderstands and misrepresents itself. Method is not, I have argued, a more or less set of procedures for reporting on a given reality. Rather it is performative. It helps to produce realities. It does not do so freely and at whim. There is a hinterlands of realities, of manifest absences and Otherness, resonances and patterns of one kind and another already being enacted and it cannot ignore these. At the same time, however, it is also creative. It makes new signals and new resonances, new manifestations and concealments, and it does so continuously.

If Law's position is adopted, or something akin to it, then the attempt to measure is at best misleading and approaches to assessment have to be rethought. Yorke (2005a) points out that such thinking puts judgement in place of measurement. This is a theme to which we return in section 8, below.

2. Fuzzy achievements

The achievements valued in higher education and by employers, of which the nine described above are a sample, tend to be complex, to be fuzzy (see section 1, above). This fuzziness is associated with:

- a. Lack of agreement about the nature of what is being assessed. Is it a skill, a quality, competence (or competency), attitude, understanding or disposition? Different answers point towards different assessment practices. 'Competence' is used in this paper but in a loose and non-committal way.
- b. Lack of agreement about what the 'competence' includes and excludes. We have seen that even something as commonplace as listening is deeply problematic – is it an internal process or an internal and external process? Where are the boundaries between listening, reading, communicating and conversing? Recall that Rhodes and colleagues (1990: 66) remark that 'it is impossible to assess a construct that cannot adequately be defined'.
- c. Style of definition. We have seen a variety of ways of trying to define a competence. There is the mentioning strategy, as if the term were so obvious that definition would be superfluous; there are attempts to produce definitions; there are extended descriptions of what competent people should be able to do; there are allusions to situations in which competence should be manifest; and in one or two cases, there are simply references to activities that should foster competence. In other words, criteria take a range of forms, some of which are potentially more useful for assessment purposes than others. Rarely do criteria meet the standards that would be expected in 'normal' social science research.
- d. Granularity. In focussing on 'listening' are we working at too fine a grain – would it be better to treat it as part and parcel of conversing or communicating?

3. Cost and convenience

- a. Logistics. How feasible is a particular assessment approach in a typical university programme in subject area?
- b. Expense. An assessment approach may be feasible but is it affordable in a typical university programme in subject area? Bear in mind that there are development, administration, analysis and reporting costs.

4. Reliability

- a. Reliable tests meet a number of technical criteria, which are especially rigorous when it comes to standardised tests. If a listening, EI or oral communication test is to be credible, it must meet threshold standards.
- b. More loosely, reliable assessments are ones in which there is only a small chance that extraneous factors influence the result. In practice this means reducing uncertainty about the marking scheme (by the use of rubrics, model answers or criteria); increasing certainty that marks are not flukes (by assessing the same competence repeatedly with a variety of task types); and minimising marker error (by computer-marked assessment, marker training and using more than one marker on essays and suchlike tasks).
- c. Reliability is costly, so compromises may be made. For example, it is easiest to get high reliability levels when tasks are straightforward and both delivered and graded by computer. This may be a problem when assessing wicked competences.

5. Validity

There are many forms of validity but the most basic are content and face validity. Loosely, they invite us to consider whether assessments really address the phenomenon they are intended to. They often do not. For example, tests of critical thinking often resemble intelligence or logic tests. This is hardly a valid approach to the situated practices of critical thinking that are so dependent on rich knowledge of a field, as well as on smart thinking. Validity is easily compromised by attempts to maximise reliability and to cut costs.

Consider, by way of illustration, Feltovich and colleagues' (2004: 90/91) account of the eleven dimensions that affect task difficulty:

Static vs dynamic. ... Are phenomena static and scalar, or do they possess dynamic, vector-like characteristics?
Discrete vs continuous. ... Can we describe attributes by using a few categories ...?
Separable vs interactive. Do processes occur independently or with only weak interaction, or do strong interactions and interdependence exist?
Sequential vs simultaneous. ... do multiple processes happen at the same time?
Homogenous vs heterogeneous
Single vs multiple representation. Do elements in a situation afford single or just a few interpretations ...?
Mechanism vs organicism. Are effects traceable to simple and direct causal agents ...?
Linear vs non-linear. ... Can a single line of explanation convey a concept or account for a phenomenon ...?
Universal vs conditional Do guidelines and principles hold in much the same way ... across different situations ...?
Regular vs irregular. Does a domain exhibit a high degree of regularity or typicality across cases ...?
Surface vs deep. Are important elements for understanding and guiding action delineated and apparent on the surface ...?

Our point is that 'wicked' competences will tend to the more problematic end of each dimension and assessments ought to reflect that. However, Feltovich *et al.* observe that there are:

... serious consequences when the material to be learned or understood exhibits [complexity] ... in such cases:

- Learners and practitioners tend to interpret situations as though they were characterized by simpler alternatives
- Their understandings tended to be reductive – that is they tend to simplify
- They tend to try to defend their simple understandings when confronted with facts that suggest that the situation is more complex than what they suppose
- Overcoming these defenses requires practice, experience and mental effort (p. 91)

Given competences that are inherently hard to define because variations on eleven dimensions affect task difficulty, and given tendencies of learners to 'drop a level' when faced with complex tasks, it is hard to see how valid assessments of 'wicked' competences might be designed. The tendency will either be to set tasks that incline to the less complex pole on each of the nine dimensions, thereby increasing reliability. Where truly complex tasks are set, it can be predicted that learner achievement will be under-recorded as they adopt coping strategies that lead to inferior performance on truly-valid tasks.

6. Test standardisation

Many tests are trialled and standardised with school children or with undergraduates (especially psychology undergraduates). It does not follow that they will give good readings or useful data about others – about trainee social workers or accountants, for instance. This is partly because it can be misleading to assume that what works with one group works with another but also because we cannot assume that a test of schoolchildren has the same purpose as a test of trainee nurses.

7. Performances and proxies

Many assessment experts favour 'authentic' assessment, which means judging achievements in natural, normal settings: a trainee nurse should be assessed in wards or health centres; teachers in the classroom; accountants in businesses, and so on. Authentic assessments can be hard to arrange and expensive to conduct. They are also problematic because some settings are richer than others and some are more supportive than others, which makes it hard to separate student achievement from contexts, unless there are multiple assessments in multiple settings (see reliability, above).

Proxies are widely used. Near proxies include many simulations. Far proxies include tests of knowledge, understanding, attitudes. Their predictive validity may be low (that is to say, scores on proxies may not map too well on to competence in practice settings).

There are particular problems with using understanding or questionnaire scores because (a) there is a difference between the propositional knowledge they address and the procedural knowledge that is associated with practice (b) professional practice involves recognising when understandings or procedures need to be deployed – there seems to be some kind of co-ordinating or meta-skill at work. Proxies are domesticated tasks where practice assessments are 'in the wild'. A sophisticated example of a battery of measures of physicians' performance at work is provided by Norcini (2005). Sources of (usually quantitative) evidence for three areas of judgement: Patient outcomes (including mortality, morbidity, physiological measures, clinical events, symptoms, patient satisfaction and experiences, cost effectiveness); Process of care (including levels of patient screening, delivery of routine services, disease-specific process measures; volume of services provided. Evidence may be collected from clinical practice records, administrative databases, diaries and case logs (mainly applicable to doctors in training) and observations (four issues for consideration are described). He describes threats to reliability and validity, including: patient mix and patient complexity.

Compared to Norcini's account, most higher education assessments of 'wicked' competence are thin and may be over-reliant on dubious proxies.

8. Application of criteria.

Criteria themselves may be defective, especially by (a) being insufficiently grounded in evidence (b) being too demanding (c) not including performance features that are usually invoked by markers. For example, standards for secondary teachers do not say that teachers should like being with young people and have a good sense of humour but judgements of classroom suitability are likely to use to these 'chalk face' criteria.

The existence of agreed criteria, even if they seem to be clear to initiates, does not mean that they will be used as intended by assessors. This is well documented. Creating criteria may reduce assessor unreliability, especially if assessors are well-trained and their performances monitored, but it does not eliminate it.

9. Direct and implied judgements

Assessment tasks typically call on many competences but it would be wearisome, if not impossible, to make a direct judgement of each. Acts of evaluation, for example, involve critical thinking, analysis and synthesis; triage involves professional knowledge, communication emotional intelligence, critical thinking and evaluation; conversation involves oral communication and listening. The issue is the degree to which it is desirable to have direct judgements of achievements, as opposed to saying that someone who scores highly on triage, evaluation or conversation can be assumed to be competent in the associated operations as well. How often to frequently-instantiated competences need to be directly assessed, if ever?

10. Assessment as an appropriate focus

There is a growing feeling that we make a mistake by treating assessment in isolation and a greater mistake by taking summative assessment so seriously. Boud and Falchikov's (2006) analysis is summarised in section 8 and complemented by Knight and Yorke (2005b) and Loacker (2005). Their insistence that good pedagogic practices should take priority over summative assessment requirements is consistent with the European Association for Research on Learning and Instruction's (2005) 'learning integrated assessment system', interest in formative assessment (Yorke, 2006), and in 'learning-oriented assessment' (Carless *et al.*, 2006).

Given the difficulties in assessing complex achievements, as represented by the brief comments on nine of them in six professional areas, the 'back to pedagogy' analysis may have attractions, which will be considered in Chapter 3.

Issues and competences

Table 6.1, overleaf, shows, very simply, some relationships between issues and competences. It is, of course, no more than our interpretation of the material reviewed in section 5 and it reflects the success of our literature searches. Other readings and other judgements might alter the pattern.

Table 6.1. Issues and competences

Competence Issue	a. Supportive relationships	b. Emotional intelligence	c. Group work	d. Listening	e. Oral communication	f. Professional knowledge	g. Relating to clients	h. Self-management	i. Taking onwards
1. Metaphysics		√		√		√		√	
2. Fuzziness	√	√	√	√	√	√	√	√	√
3. Practicalities				√		√		√	
4. Reliability			√	√		√			
5. Validity		√		√		√			√
6. Standardising									
7. Proxies		√		√		√		√	
8. Criteria		√	√	√					
9. Direct & implied	√								

Glossary for Table 6.1

1. *Metaphysics*. Does the competence exist? If so, how might we learn about it?
2. *Fuzziness*. How much agreement is there as to what the competence is?
3. *Practicalities*. What are the pragmatics of assessing this competence?
4. *Reliability*. Are results artefacts of the instruments, markers and assessment arrangements generally?
5. *Validity*. Do the assessments really get at the competence, or do they address simplified or partial versions of it?
6. *Standardising*. Have assessment tests been standardised on people similar to those we want to assess?
7. *Proxies*. How far should performance be assessed, as opposed to basing assessment on proxies, such as level of propositional knowledge)?
8. *Criteria*. Are there difficulties in forming clear and usable assessment criteria? In using them in practice?
9. *Direct & implied*. Is this competence to be assessed directly, or is threshold achievement to be inferred from performance on other tasks?
10. *Appropriate focus*. This is a meta-issue, covering all 'wicked' competences and is not addressed at this stage

7. Survey data

The design of an empirical investigation of the assessment of 'wicked' competences in higher education was outlined earlier in this chapter. This section reports on data from the Autumn 2006 survey, which yielded 83 respondents.

Informants confirm that these competences are, by and large, important and should be assessed:

Competence	a. Supportive relationships	b. Emotional intelligence	c. Group work	d. Listening	e. Oral communication	f. Professional knowledge	g. Relating to clients	h. Self-management	i. Taking onwards
Valued by employers etc?	Yes ⁷		Yes	Yes	Yes	Yes	Yes	Yes	Yes
An assessment priority?	Yes		Yes	Yes		Yes		Yes	

It will be seen that there are three cases where it is reported that 'wicked' competences are important but that they lose out to competition for assessment time from other competences.

This is mildly consistent with our informal hypothesis that data would show the assessment of 'wicked' competences to be deeply problematic. There were three grounds for this expectation. One was the mismatch between commonplace assessment practice and the standards of social science enquiry. A second was the ramshackle nature of definitions of these competences in the six subject areas. Given the tenet that ill-defined outcomes resist assessment, especially summative, measurement-style assessment, it was expected that these competences would be causing higher education practitioners real assessment difficulties.

The third reason for our hypothesis was recent work pointing to endemic problems in university assessment practices (Boud and Falchikov, 2006), with Yorke (2005a) identifying particular problems in the assessment of work- or practice-based learning, where 'wicked' competences might be supposed to be prominent.

⁷ Many of the items on the questionnaire asked for responses on a five point scale. There is a temptation to give these numbers a status that they cannot have and begin to carry out complex parametric procedures on them. Instead, we grouped them into three sets: one set comprises responses clustered around the mean and ranging from 2.5 to 3.49. A second set is those in the range 3.5-5.--, with the third being those in the range 1.0-3.49. We then show only the positive or negative responses. Blank cells indicate that responses clustered on the middle ground.

Some support for this hypothesis was found in responses to question 2, which asked whether assessment of these competences was mainly at module or at programme level.

	Accountancy (n = 19)	Early years teaching (n = 4)	Nursing (n = 20)	Secondary School teaching (n = 10)	Social work (n = 18)	Youth work (n = 12)	All (n = 83)
Mainly programme / award	3	1	3	5	3	3	18
Mainly module, course or unit	16	3	17	5	15	9	65

Theory (Claxton, 1998) argues that these competences are slow-growing and leads to the conclusion that programme-level assessment, run over at least three years, would be most appropriate.

However, in other respects, questionnaire data from 83 respondents does not fit our hypothesis very well. In the following sections, data are summarised by competence under three headings:

- a. The priority given to assessing 'wicked' competences
- b. The degree to which such assessments cause difficulties
- c. The methods used to assess these competences.

In looking at the degree of difficulty attaching to the assessment of a competence, we decided to highlight scores of 3.5 and above (on a five point scale, where 1 indicates strong agreement that something was *not* a difficulty in the assessment of a named competence). Scores lower than 2.5 were also highlighted to show where difficulty was really perceived. The surprise was the lack of difficulties and the number of reports of non-difficulty.

Data on priorities are not remarkable and descriptions of the range of assessment approaches are interesting only in a purely descriptive way.

a. Developing supportive relationships

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Secondary School (n=10) and Youth Work (n=10)	Developing supportive relationships	3.6	3.6	3.6	3.1	3.6	2.6	Not a problem
Secondary School (n=10)	Developing supportive relationships	3.5	3.7	4.0	2.9	3.3	2.8	Not a problem for Secondary education
Youth Work (n=10)	Developing supportive relationships	3.7	3.5	3.1	3.3	3.8	2.4	Not a problem, yet Q24 ⁸ says may be a problem for Youth work

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Supportive relationships (n=20)	2.9	2.6	3.35	2.5	4.15	3.25	4.3	3.45

Assessment by tests and coursework predominate. It is arguable that they are not fit for purpose. There seems to be some uncertainty as to whether these competences are assessed directly or indirectly. There seems to be muted interest in other assessment approaches.

⁸ Q24. 'There is a worry about the reliability of assessments of this competence.'

b. Emotional intelligence

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Youth Work (n=2)	Emotional intelligence	1.5	1.5	2.0	2.3	2.5	1.5	Problematic

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Emotional intelligence (n=2)	3.5	0.5	4.5	1.5	0.5	3.0	1.5	3.0

Insofar as there is a tendency, it is to direct assessment of emotional intelligence. It is impossible to see how this direct assessment is done, although there is some recognition of the potential of portfolios and coursework. Other methods, arguably with equal or greater potential, are disdained.

The surprise is that clear criteria exist – perhaps this is an artefact of a sample of two, though.

c. Group work

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Accounting (n=15)	Group Work	3.1	3.0	3.1	3.0	2.8	2.4	Neutral, bordering on not problematic, yet Q24 ⁹ says may be a problem

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Group work (n=15)	3.07	2.8	3.13	3.8	3.07	3.87	3.53	1.87

Again, there is some equivocation about the use of indirect and direct assessment approaches; so too about the existence of clear criteria. There are modest tendencies to use tests, portfolios and evidence of workplace performance. Interestingly, coursework assessment does not address this competence, implying that individual, written or on-line tasks predominate.

⁹ Q24. 'There is a worry about the reliability of assessments of this competence.'

d. Listening and assimilating

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Nursing (n=8)	Listening and assimilating	3.1	4.1	3.9	3.0	3.5	2.5	Not a problem

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Listening (n=8)	2.25	3.13	3.13	2.38	3.38	2.88	3.5	3.13

Responses generally cluster around the mean, with the only strong exceptions being the view that listening is neither directly assessed nor assessed through workplace performance (which might have been thought to be the most valid approach). One inference might be that there is uncertainty or vagueness about assessment methods in use.

e. Oral communication

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Accounting (n=4) and Early Years (n=1)	Oral Communication	2.6	3.4	2.4	2.7	2.8	1.6	Neutral, bordering on problematic, Q24 ¹⁰ says is a problem
Accounting (n=4)	Oral Communication	2.3	3.0	2.5	2.5	2.5	1.5	Neutral, bordering on problematic, Q24 says is a problem for Accounting
Early Years (n=1)	Oral Communication	4.0	5.0	2.0	3.3	4.0	2.0	Not a problem except for time, Q24 says is a problem for Early years

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Oral communication (n=5)	2.4	2.4	2.62	3.8	3.6	4.4	4.2	2.85

There is uncertainty as to whether oral communication is directly or indirectly assessed but several assessment methods are reported. Notice the low-ish scores for the existence of clear criteria and for coursework assessment, though.

¹⁰ Q24. 'There is a worry about the reliability of assessments of this competence.'

f. Professional subject knowledge

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Social Work (n=13)	Professional Subject Knowledge	3.6	3.9	3.5	3.3	3.3	2.7	Not a problem

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Professional subject knowledge (n=13)	1.85	3.15	2.62	3.46	3.69	3.38	3.54	2.85

Clearly – perhaps surprisingly – subject knowledge is not directly assessed but there is limited enthusiasm for saying that it is therefore indirectly assessed.

There is a lack of clarity about criteria and methods used to assess it, except for a distinctive emphasis on simulations and some propensity to use written tests.

g. Relating to clients

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Nursing (n=12)	Relating to clients	3.2	3.9	3.0	2.6	3.3	2.2	Not a problem, yet Q24 ¹¹ says may be a problem

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Relating to clients (n=12)	2.17	2.00	3.5	1.92	3.58	3.0	3.42	3.5

Assessment of this competence looks problematic. There is no view about whether it is directly or indirectly assessed. Workplace assessment, which might have been thought the most valid, is not used. There is a slight tendency to see criteria as clear and to use simulations and coursework. Unfortunately, this questionnaire cannot shed light on ways in which these methods provide performance data on this competence.

¹¹ Q24. 'There is a worry about the reliability of assessments of this competence.'

h. Self-management (confidence and effectiveness)

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Early Years (n=3)	Self management - confidence and effectiveness	4.0	3.3	3.7	3.6	3.7	3.0	Not a problem

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
Self-management (n=3)	2.0	2.67	2.33	2.67	4.0	3.0	3.8	3.6

Simulations, tests and coursework were all cited as primary assessment methods, although it is hard to see how tests address confidence and effectiveness. Vagueness about the use of direct as against indirect approaches is complemented by problems with the criteria that are available.

i. 'Taking it onwards' – acting on diagnoses

Difficulty

Subject area	Competence	Training	Cost	Time	Criteria	Meaning	Reliability	Conclusion
Social Work (n=5)	Taking it onwards' - acting on diagnoses	3.0	3.4	4.0	2.5	3.2	2.0	Not a problem, yet Q24 says may be a problem

Blue means 'not a problem'; Red means 'problematic'

1 = Strongly agree that there is a problem here; 2 = Agree; 3 = Neutral; 4 = Disagree; 5 = Strongly disagree

Assessment methods in use

	WCs directly assessed (q17)	WCs indirectly assessed (q14)	Clear criteria exist (q21)	Assessment mainly on workplace performance (q13)	Assessment mainly on simulations (q15)	Assessment mainly on portfolios (q19)	Assessment mainly on tests (q20)	Assessment mainly on coursework (q23)
'Taking it onwards' (n=5)	2.0	2.8	3.4	2.0	4.0	3.0	3.8	3.6

The lack of attention to assessing this in the workplace is surprising, although the lack of clarity about the relationship between direct and indirect assessment is not, at this stage of the report. It may also surprise some that such an esoteric competence – we have certainly never seen it cited elsewhere – is associated, albeit gently, with the existence of clear criteria.

Discussion

It has been shown that survey data did not bear out the informal hypothesis that the assessment of 'wicked' competences would be seen as particularly difficult. One explanation is that there really is not a problem here. However, there are alternatives. Recall that the hypothesis was based on deliberations on social science method, the quality of statements about the competences in the subject areas and some research indicating marked problems with assessment in higher education.

Two ways of explaining questionnaire findings are:

Inappropriate research methods. One unsolicited report on the questionnaire aid that it had stimulated so much thinking that response had proved impossible. The issues it raised needed extended consideration with others. Consequently, it would have been inappropriate to answer it. Questionnaires are notorious for forcing complex understandings into simple response formats, with consequent validity problems (Oppenheim, 1992). Our unexpected results may be an artefact of the enquiry method.

False consciousness. It is possible that respondents identified so few difficulties because they have not thought deeply about enquiry methods and issues in the assessment of non-determinate and complex outcomes. In short, they might be deluding themselves into reporting few difficulties where, to an expert, there is a plethora of them. This explanation has the advantage of fitting with the foregoing analysis of assessment, method, and 'wicked' competences and the massive disadvantage of casting teachers in higher education as dupes or dopes.

Phone interviews with a sub-sample of respondents were done in December 2006 to get a sense of which, if either, of these explanations was more sustainable.

8. Telephone interviews

The project plan allowed for a dozen or so telephone interviews to be carried out once the survey data had been analysed. There was interest in seeing whether the interview data tended to support any of the three provisional hypotheses that had been used to explain the survey findings (there is not a problem assessing 'wicked' competences; the unexpected results are attributable to inappropriate enquiry methods; there is a problem but it is cloaked by teachers' 'false consciousness'). The interviews were organised around four prompts:

1. Is it problematic to assess the oral communication skills of Accounting/Nursing/Social Work students?
2. If yes, what do you see are the difficulties associated with assessing these skills?
3. If yes, can you suggest any solutions?
4. Do you trust the judgements of Accounting students' oral communication skills that are made by others – in their departments, in other universities and in Accountancy practice?

The first three prompts invite informants to revisit, in rather general terms, territory that was more closely explored in the questionnaire. In the event, informants tended to identify problems that had not been highlighted by the questionnaire, which suggests that the questionnaire method was not ideal for uncovering problems¹². The fourth question is designed to force informants to consider assessment problems that they might have tended to overlook. The underlying thinking with this prompt was that informants might feel that they had the assessment of 'wicked' competences under control in their own modules or units and be replying to us accordingly. Prompt 4 shifts attention to the trustworthiness of a student's competences across a programme of study as a whole¹³. Responses to this prompt tended to indicate that there are, indeed, hidden assessment problems and that the false consciousness hypothesis is sustainable.

Informants had volunteered to be interviewed and were drawn from the three best-represented subject areas 14. Interviews were done in December 2006 and early January 2007. The interviewer kept headline notes and recorded the conversation. Recordings were fully transcribed and Nvivo was used in data analysis. Data are reported, in summary form, by subject area.

¹² A nice point is that we cannot assume that the questionnaire method was defective. It may be telling the truth and the interviews might be misleading. It is possible that the telephone interviews picked up those informants most aware of the assessment issues and that the very fact of being interviewed suggested to them that we were still looking for 'the right answer'.

Another interpretation of the questionnaire-interview difference is more in line with Law's (2004) reasoning. It says that both the 'few problems' pattern described by questionnaires and the 'more problems' described by interviews are both true. Sometimes one pattern prevails, other times the other.

¹³ Mantz Yorke reports that, un-prompted, academic staff often fail to appreciate just how problematic the English practice of degree classification is. Prompt 4 assumes that the same may be true with the overall assessment of 'wicked' competences.

Analysis of Accounting interviews

Respondents to the questionnaire seemed to agree that oral communication is problematic to assess (particularly about whether colleagues are trained to assess it and how much time is available to assess it) and seem less concerned about Group work. For oral communication the key informants had suggested that:

- There is always the risk of cultural and perception difficulties in assessment of these skills;
- It is important to be able to explain accounting techniques/terms/decisions clearly to non-accountants.

Five people from Accounting (all of whom teach at University level), were specifically asked about assessing oral communication skills of accounting students. The following problems were highlighted:

- Accountants are not generally trained to assess oral communication skills;
- The importance of oral communication is crowded out by the amount of information about accounting that students have to learn;
- Good oral communication skills are also less valued by the students because if they are assessed at all, the marks count for only a tiny percentage of their final course or module score;
- The traditional assessment methods of examination leading to professional qualification are still what is provided, there is little done to assess the soft skills required, though some universities are tackling this by working on presentation skills with their students, and these are the universities which have a good reputation for turning out good graduates who are more likely to get the better jobs;
- In presentation sessions accounting students are explaining to accounting students or tutors rather than those who are not familiar with the terminology, so they get little real practice on how to explain terminology or accounting decisions to lay people;
- Post graduate courses for accountants tend to be refresher courses or updates on latest tax or financial issues, and do little to emphasise the need to build oral communication skills;
- There is a mismatch in perception between what the students believe they will be learning in an accounting programme, and what the profession claims it wants or needs students to learn – students think they are going to learn about good financial record keeping, whereas the modern profession emphasizes communication of financial information as being important;
- It would require some staff development of those doing the assessment of students to improve the assessment of oral communication skills;
- Developing a student's ability to think critically and logically is usually done through their writing skills in the first instance (because of the amount of information that needs to be taught on the courses), even though arguably oral communication demands a higher level of critical thinking and ability to explain than writing, because in a question and answer session there is no time to redraft a reply;
- People have individual styles, and cultural and personal background also often inhibits understanding.

In some of the conversations, the methods of assessing presentation skills were discussed at some length, as it is in making presentations to other students and assessors where the oral communication skills are judged, and then only in an incidental way. It was generally agreed that some sort of proof or artefact of the student's performance needed to be produced for their assessment, and for oral communication or presentation skills the obvious one would be a video recording, which could also be used as a teaching tool to show students good and bad points of their presentation, and to help other students see good presentation skills in action. But students

were reluctant to be filmed, so usually had to provide written reports summarising their presentation as well as their OHPs or Powerpoint slides, but this of course would not show the actual performance. Two people interviewed saw the value of presentations as a way of teaching students and assessing their understanding of the course materials, as mistakes in an examination could not be queried, but a slip up in an oral presentation could be raised and discussed at the time; however this method of assessment is very expensive in staff time. One interviewee observed that some students who might actually be quite good oral communicators on a one to one basis were reticent about speaking up in class in front of others.

The following solutions were suggested in the course of conversation by the various interviewees, and some of them were already being explored or tried in some universities:

- Provide Peer and Self assessment sessions in which students are given the opportunity, in groups, to decide what needs to be assessed in their presentations and how it should be marked (negotiated assessment);
- Provide post-graduate training which is geared to developing oral communication and group work skills;
- Include 'making presentations' in the undergraduate curriculum;
- Change staff attitudes towards the importance of assessing these skills;
- Teach students to work out what the client wants to know, and not just to concentrate on the figures but what can be read from those figures or how they can be used, as this is what will be required in the job market (analytical skills);
- Use a series of different approaches to improve their communication skills, including utilising video;
- Use the practicing of oral communication skills as a means to improve their critical thinking and analysis abilities;
- Phase in these improvements from the first year of an Accounting degree programme in stages.

The trustworthiness of others' judgements of competence was assumed to be assured by means of having more than one assessor and keeping audio, video or PowerPoint records of presentations. One informant mentioned that 'the team' is stable and so there 'is an element of trusting each other's judgement ... and surprisingly on most occasions the average mark and range is acceptable in statistical terms'. The overall standards and framework provided by the Chartered institute of Public Finances and Accountants was thought to help the profession to have shared understandings and standards, which are operationalised by,

traditional British exams with some course work thrown in but the coursework tends to be marked internally and academics can't always be trusted, so the final thing is the [CIPFA] exam. So if you are putting on a course for students leading to externally-sat final exams ... and if those exams are all written, at best you can only describe how they would do something orally Unless the oral thing helps them learning towards those exams, then forget it really.

Another informant looked at presentations within coursework and was rather cautious: in accounting colleagues will feel a little bit uncomfortable in assessing that sort of soft skill ... Top end is very easy [and] you can certainly distinguish between 60-something and 40-something ... but you might not be able to distinguish between 61 and 65.

Generally, I do [trust others' judgements], within broad bands ... In any one grade there's going to be an element of subjectivity but that is ... true in any exam, so I don't really think that's the end of the world.

In one case pragmatism prevailed over unease:

The short answer is 'yes I do [trust others' judgements]'. Despite all the reservations I've alluded to earlier. Yes I do but only for pragmatic reasons ... frankly there isn't an alternative.

Notable was a comment by the informant who referred to the development of broadly-shared understandings by a stable team. First, a problem was identified with 'communication apprehension' where the team knew that 'to cope with that properly is a very expensive process ... and the university isn't in a position to provide that resource so we've got to do what we can under the circumstances'. This led to more general comment:

As far as the profession is concerned, I think I would say I am sceptical [about the assessment of oral communication] ... Some research that we've done and that's published shows that levels of communication apprehension really don't move through accounting degrees and professional development ... what employers see as the priorities [in the first degree] and after the training period in industry they still saw that oral and written communication skills were still the priority. So, in a sense, what they were saying is "you're not doing the job". So I think the profession needs to rethink ... there's a huge problem in its processes ... and in its assessment process ... in a sense the profession is saying "the key skill is communication, so where are you assessing it?" And the answer is that they're not.

Finally, one said 'I don't know whether the qualifications are at levels that we expect involve oral communication skills or not ... But if they do, I wouldn't trust it or it's certainly weaved in at an inappropriate level in my view'. This informant added that it was easy to identify students who 'are at ease with themselves' and therefore good communicators. More could be done to help others 'given that these things are terribly important to them' but adding that 'the one thing they hate the most is the presentations'.

Analysis of Nursing interviews

Respondents to the questionnaire did not seem too concerned about Listening and Assimilating skills, but seemed less sure about how to assess whether a student nurse is good at relating to clients. For relating to clients the key informants had suggested that:

- Nurses should provide care which demonstrates sensitivity to the diversity of patients and clients;
- Be able to give clear explanations, empathise with patients, access and respond to patients' reactions to their treatments/the information being given;
- Display the ability to get on with people – build a collaborative partnership with patients.

Four people from Nursing (all of whom teach at University level) were specifically asked about assessing the 'relating to clients' skills of nursing students. The following problems were highlighted:

- Assessment of this skill is often done indirectly through assignments or in clinical laboratory setting rather than in a practice based setting;
- When assessing these skills in real life setting of clinical practice, it is hard for the assessor to be objective;
- Assessors are happy to assess technical skills but are less comfortable with assessing personal skills;
- Current assessment methods of this skill are not comprehensive and are not always clearly articulated to assessors;
- Students are still being passed (if they pass the academic side of their course) even if there is some doubt about their ability or understanding of how to empathise with patients – a failure to fail them;
- Students in the full degree programme have less opportunity to get involved with patient care and feel more removed from patients than student nurses taught before the degree programme was brought in;
- Inexperienced nurses are sometimes the ones assessing or mentoring the student nurses, and sometimes these are poor role models for the student nurses to follow;
- The level of expected ability is not sufficiently taken into account when assessing the student, resulting in students being penalised unnecessarily;
- Assessors in practice situations are too busy to spend time analysing the competencies to work out how best to assess them;
- Students nurses may not always recognise when they have been effective in a caring relationship with a client, and their efforts are not always acknowledged appropriately;
- A student's own cultural or social background and life experiences may result in inappropriate or misunderstood interactions with patients.

Students are placed in a variety of different clinical areas for practice-based learning, so they encounter a wide variety of approaches to care. This is good, but needs to be managed carefully so that they can get the best guidance of how to handle the different situations they encounter. Therefore it is crucial that problems with their mentor and assessor relationships need to be identified early in a placement. Problematic attitudes towards clients can be displayed in their written work as well as in their practice, so it is not inappropriate to assess these skills indirectly as well as directly.

The following solutions were suggested in the course of conversation by the various interviewees:

- Teach students to use reflective practice and take personal responsibility for their actions;

- Provide more opportunities for student nurses to interact with patients in a variety of supported placements;
- Assess students in 'real life' situations by observing behaviour and provide clear feedback on their performance;
- Provide mentors and assessors with clearer guidance and checklists on how to assess these skills, and make them aware of the things that might influence a student's behaviour (for example cultural background);
- Encourage assessors to be more rigorous in their approach to marking students in practice-based situations, and not just pass them because it is too hard to make a definitive judgement;
- Choose practice-based mentors and assessors carefully, not conveniently;
- Require assessors to provide evidence to support the assessments they make;
- Write competencies in more user friendly language and give students specific examples that illustrate the competency of relating to clients (Refer to the NMC Code of conduct and the NMC's Outcomes, Standards for Proficiency Registration for specific examples of attitude and behaviour skills). Get tutors, assessors or student nurses to work in groups to analyse these examples as a means of learning from them and how to use them to assess these skills;
- Follow up discipline issues (that cause problems with clients) which arise from a student's attitude, make sure that once such issues are dealt with, the appropriate debriefing takes place and support is provided to the student for improvement.
- Praise good examples of a student's ability to relate well to clients, as students are more likely to repeat good behaviour if it is recognised and acknowledged.

Responses to questions about the trustworthiness of others' judgements of competence were mixed, as this informant's response shows:

No [I don't trust them] ... I suppose in general, yes [I do trust them] but I suppose within nursing there has been quite a lot written about the fact that there quite a lot of evaluations of nursing programmes have shown that students don't feel fit to practise when they qualify ... I think in terms of relating to clients the clients are less likely to, my feeling is, they are less likely to fail the students because of those kinds of issues. It's much easier to fail the student if they [clients] know they don't have a particular technical skill. But it's much more difficult to, I suppose, quantify and rationalise why you would fail someone in the seeming lack of relationship.

More forthright was another informant:

No I don't [trust others' judgements]. I'm really difficult – no I don't. But in the real world we have to ... But I believe assessment is subjective anyway. The problem I've mainly encountered is ... inexperienced nurses assessing students ... the other side of the coin is that we have over-zealous assessors who crucify students ... until we can articulate what it is we want nurses to do and think and behave, we're not assessing them at all.

An interestingly balanced response was:

I do [trust others' judgements] if they provide the supporting evidence ... [and] if there's robustness to the evidence so that it's not just a one-off situation.

I feel that practitioners who are supporting students learning in practice are becoming more confident and sophisticated in assessing students with regard to attitude and behaviours ... because it's been made more explicit to us... [and] I think we're much more aware of our responsibilities.

Analysis of Social Work interviews

Respondents to the questionnaire seemed to agree that 'taking it onwards – acting on a diagnosis' is problematic to assess and seem less concerned about how to assess Professional Subject knowledge (though this is a complex area for Social Work students to absorb). For 'Taking it onwards – acting on a diagnosis' the key informants had suggested that:

- Social workers should be able to inform individuals, families, carers, groups and communities about their own and the organisation's duties and responsibilities;
- Develop participative approaches to identifying needs and then making plans to respond to those needs.

Five people from Social Work (all of whom teach at University level) were specifically asked about assessing 'taking it onwards – acting on diagnosis' skills of social work students. The following problems were highlighted:

- Diagnosis is not a good word to describe this competence, as social workers don't make diagnoses, rather they assess a situation based on several factors and make judgements and recommendations as to what should be done to improve or resolve that situation;
- Even the best students overdo their judgements of a situation by trying to link the theories they have learned to a scenario or a real life situation when they don't have enough information or evidence to make such connections;
- Social work practice is very individual and subjective in nature, each case is different and context driven;
- Social work students are often assessed in a practice environment, with assessments carried out by a practitioner, who is often employed in that practice area rather than by the University running the degree programme, and therefore there is a level of trust involved, and an acceptance that mistakes may be made;
- The judgements made by the assessors are often personal and individualistic, based on a particular time, place and case, it is difficult for them to reach completely objective assessments of students' judgements of a situation, though they may have specific institutional criteria in mind;
- Assessors in the practice setting may have difficulty in knowing at what level to assess the student and may not have sufficient awareness or guidelines from the university programme as to how to assess the student (each university programme varies);
- The QAA benchmark states that Social work is a moral activity, which indicates its complexity from the very start;
- A student needs to cope with different sets of criteria, regulations or procedures to follow depending on where their placement takes place – it may be a statutory or non-statutory role.

Students need to learn to manage a situation in an engaging way which does not oppress those they are trying to help, and this depends in a great part in how they communicate, engage and share the decision making process with those who need but may not wish to be helped. Students need to be able to deal with the complexity of emotions, regulations, conflicting demands and explanations and negotiate solutions or plans of action quickly and these skills take time and experience to learn. Assessment is also based on student portfolios of the records they keep of the work that they do in a practice based environment, these portfolios are used as a means of teaching a student how to analyse a situation as well as assessing what they have learned.

The following solutions were suggested in the course of conversation by the various interviewees:

- Arrange for practice to be directly observed by an assessor, preferably more than one assessor (though having more than one assessor is not always practical or possible);
- Practice-based assessors should be provided with clear guidelines about the level at which their assessments of a student's ability should be made for that particular University programme;
- Group discussions should be held between students and their university based tutors about the safety of the claims they make based on their academic (theoretical) knowledge to inform their decision making about a particular case or scenario;
- Teach them to self assess their judgement and decision making process by providing them with guidelines of how to question their own practice, and the opportunity to discuss this self assessment with a knowledgeable mentor;
- Provide students with the 'right kind' and variety of practice opportunities;
- Encourage students to use evidence of following a process (decision making models) as part of the criteria to show how well they are coping with the complexities and uncertainties when reaching a judgement/making a case and acting upon it;
- Acknowledge and take account of the individuals and organisational context in the assessment process, including holding placement agreement meetings, mid-point reviews and end of placement reviews, strengthening the placement as an assessment based learning opportunity.

Asked about trusting others' judgements of wicked competences informants normally appealed to the expertise of the practice teacher who supervises the student, the evidence of the portfolio and the arrangements for second and, if need be, third readings of the evidence. This informant also saw discussion with the student as a valuable element. This was quite a common theme in conversation about the trustworthiness of others' judgements. Two points are notable. First, the arrangements are very expensive, especially when 'what we would use generally is direct observation for that in terms of hired work'. Secondly, there is a deep assumption in these five responses that there are shared understandings of the competence, standards and criteria. This confidence is, of course, necessary for the programme to function. It does rather fly in the face of academic work that finds multiple and inconsistent understandings.

There are some direct comments on multiple meanings, standards and criteria. One informant said:

Well, I don't think you've got any choice, really. I think you have to and there is a constant dilemma in Social Work about the key decisions made via people over whom the university has no QA control ... we do work on the assumption that people have arrived at decisions in good faith ... we're never going to be in a position to gainsay their judgement anyway.

Another said 'all these judgements are essentially negotiation ... on occasion when we've looked at the evidence of portfolios we've been led to ask "is this safe practice?", even if a team's passed it'. A third said, 'I don't think we have quality assurance mechanisms to sort of check up on their judgements but you don't have any option but to [trust them]. If we didn't trust them we couldn't have a programme.

The three hypotheses

There is no great problem assessing 'wicked' competences.

The questionnaires identified fewer problems than expected. However, the interviews identified problems and offered solutions. Scientifically speaking the null hypothesis – there is no great problem – should be retained unless there is compelling evidence to the contrary. The interview data should probably not be seen as sufficiently compelling to reject the null hypothesis. A judicious compromise might be offered: unprompted, practitioners identify fewer assessment problems than might have been expected. Prompted, as in the interviews, they describe problems, especially when their attention is directed away from their own local assessment practices to the overall, programme-wide assessment of competence.

Unexpected survey results are attributable to the use of inappropriate enquiry methods

Certainly there is a difference of emphasis in the findings from the interviews and the findings from the questionnaires. However, it may be naïve to explain this in terms of a method being inappropriate. It might be that different methods and questions instantiate different aspects of informants' thinking and experiences. This second explanation follows Knight (2001) and Law (2004). They say that conventional accounts of cognition assume that people have coherent concepts and that, on this view, the task of science is to identify and describe it. They argue that, contrariwise, cognition is more fractured, less coherent and more situationally-specific. Science can describe its folds, incoherences and gaps but should not try to represent inherent mess by a singular and contradiction-free account.

If this second view of social science is plausible, then different methods have, predictably, led to different representations of assessment practice. If the first view is preferred, then further enquiries are called for to determine the truth of the matter.

Unexpected survey responses are attributable to teachers' 'false consciousness'

The interviews provide some support for the view that, unprompted, informants seem to think that practice – by which they probably mean their own local practice – is working well enough and that it is not a source of disgruntlement. However, when this consciousness is challenged, notably by the fourth interview prompt, problems are disclosed. This is taken to be some support for the third hypothesis.

Other interview findings

It is notable that informants are rich in ideas for improving the assessment of 'wicked' competences, which lends some support to the view that this area of assessment is problematic.

There are also some indications, notably in accountancy, that the assessment of a 'wicked' competence valued by employers is effectively marginal in the curriculum.

It is also notable that the solutions might be called pragmatic or technical, allowing practice at the level of the individual module or unit to be improved, assuming the availability of time, training and other resources. What is strikingly absent is attention to the problem of describing student competence across the programme. It may be interesting to know that the judgement of assessors is that a student showed communicative competence in a particular course. However, this does not say anything about communicative competence in general. Programme-level data are absent, as are accounts of the ways in which this competence is understood and assessed in the programme. As was remarked in section 3.1, such information is necessary before we can trust that the judgement is fair, careful and representative of the students' communicative competence in general. It is also needed if we are to appreciate what is understood by

'communicative competence': section 3.5 showed the raggedness of many accounts of 'wicked' competences.

9. A summary of the empirical enquiry into the assessment of 'wicked' competences

There were grounds for supposing that the assessment of 'wicked' competences is problematic. Key informants in six subject areas identified 'wicked' competences that they reckoned to pose particular assessment difficulties. These were competences nominated by the key informants and not all of them were ones that could be found in professional standards or subject benchmark statements.

A survey was constructed to explore the assessment of the nine nominated competences. Networks and mail lists were used to invite participation in the survey and a total of 83 responses were obtained in Autumn 2006. This represents a very low response rate and respondents were, by definition, not representative of the population of teachers in these six areas. Nevertheless, it was supposed that their views would help to define the range of assessment experiences in this field.

In the event, survey data produced the unexpected finding that respondents saw few acute problems in assessing 'wicked' competences. Three provisional hypotheses were explored through telephone interviews with a subset of 14 volunteers from the sample of 83. Interview data tended to identify problems in the assessment of 'wicked' competences. A number of pragmatic responses and practices were described and might usefully inform work in other professional programmes. It was remarked by some informants that attending to these problems would involve giving them greater priority in the curriculum, investing in staff training, and providing more resources to allow more time to be spent.

We have observed the tendency to describe these assessment issues as course or unit issues and suggested that they need to be seen as programme issues. However, actions might be taken at many levels, ideally within a general consideration of assessment practices. That is the topic of the final section of this report.

Chapter 3 A pragmatic approach to the assessment of 'wicked' competences

Further investigations

The empirical enquiries reported here rested on the assumption that assessment problems would be known and reported by respondents. This may not actually be the case; provisional hypothesis 3, that there is an assessment problem which is cloaked by teachers' 'false consciousness', assumes that it is not the case. However, any investigation of the 'false consciousness' explanation would involve more intensive, case-based enquiries. This, of course, brings to the fore hypothesis 2, that unexpected data are the outcome of inappropriate enquiry methods. We have offered an alternative reading of this line of explanation by arguing that thinking is not coherent and that different research methods may therefore tend to bring out different – and sometimes contradictory – faces of the phenomenon in question. That only increases the interest in conducting intensive, case-based enquiries into practice¹⁴.

So, in order to understand assessment practices and associated logic-in-use it would be helpful to look 'close-up'. This might mean looking closely at a sample of undergraduate degree programmes in specified professional subject areas, concentrating on a few 'wicked' competences and examining, amongst other things:

At the module level

1. Relevant, planned assessment activities
2. The reasoning behind the design of these activities and in their implementation
3. Relevant assessment criteria, assessor comments and feedback and levels of agreement between first and subsequent markers
4. Means of reporting achievement in terms of these target outcomes

At programme level

5. The pattern of assessment of target competences across the programme
6. Levels of consistency, continuity and progression. (Are programme-wide criteria used? Is there regular assessment of these competences across the programme? Is there progression, so that level 3 assessment is markedly different from level 1 assessment of the same competence?)
7. Cumulation of assessment data in the awards process; programme-level recording and reporting of achievement.
8. Teachers' thinking about the relationship between the assessment of 'wicked' competences at module and programme levels.

Ideally, this would be complemented by enquiries involving students, employers and professional associations. We recognise three difficulties here:

- Knight's experience in another university suggests that students may tend to over-rate their command of 'wicked' competencies, especially where there is no consistent and prominent programme-level attention to them. Ironically, the more a programme takes 'wicked' competencies seriously, the less highly level 1 and 2 students rate themselves.
- Studies of employers' thinking about student employability (for example, Knight and Yorke, 2004) report difficulties engaging employers and reservations about the validity of their comments.
- Professional associations are often the sponsors of standards that may themselves be deficient (see pp. 18-32, above). They may have limited goodwill towards enquiries that may highlight problems.

¹⁴ Explanation 1 has been discounted at this stage because there are many reasons to suppose that there *ought* to be problems here. However, this should really be treated as a provisional suspension of explanation 1, not as a scientific certainty.

This set of enquiries would outstrip the Practice-based Professional Learning Centre's funding.

Yet – and notwithstanding the questionnaire study reported above – there are grounds for supposing that all may not be well with the assessment of 'wicked' competences, not least because the assessment of achievement is, in general, a matter of concern in higher education, as the protracted deliberations of Burgess's committees on judging, recording and reporting achievement show. Suggestions for enhanced assessment practice – and not just for better assessment of 'wicked' competences -- are offered under four headings. The claim is that, insofar as there are problems in assessing 'wicked' competences – and we recognise that there is some uncertainty here – then these general suggestions can offer some relief.

Suggestions for improving assessment practices

Before offering sets of suggestions, it is important to acknowledge that practitioners are not convinced, according to the questionnaire data, that there are serious problems in assessing 'wicked' competences. Alongside this reaction, there are interview data and the tenor of writing on social science and assessment in general, which all indicate that there are – and ought to be – significant failings in assessment practice in this area. Similar reactions have been observed in public responses to the Burgess Committee's proposal that degree classification should be abolished in the UK: 'where's the problem?' say some commentators, even though classification practices are a scientific abomination.

We consider that any attempt to improve practice in this area needs to begin with the assumption that many, perhaps most, teachers will fail to see that there is anything especially problematic about their pragmatic solutions to the challenge of assessing 'wicked' competences. This is a double challenge for innovators. Not only does assessment practice have to be improved, but colleagues need to be convinced of the need to improve it in the first place. For this reason we have chosen to couch our suggestions for enhancement in terms of general specifications for sound assessment practices.

These suggestions are based on an understanding of the assessment literature and on discussions with educational developers in UK, Hong Kong, Belgium and Austria, as well as on conversations with assessment scholars outside the UK. They are couched in terms of sets of expectations for different actors in higher education. Adoption – even partial adoption -- should enhance assessment practice in general, in the process assuaging any problems experienced in the assessment of 'wicked' competences.

Teachers

An enlightened practitioner might know about most, if not all, of these themes and have some commitment to working them out in everyday practice.

Topic	Scope
Variety of assessment tasks	Individuals should know the range of assessment tasks in general use and those commonly used in their subject area. They will understand that different task formats favour the assessment of particular learning outcomes. They will appreciate advantages and disadvantages that are often associated with different forms.
Assessing the range of learning achievements	Individuals will know ways of assessing more complex learning outcomes, including 'soft skills', as well as of assessing more determinate outcomes (knowledge, application of heuristics, performance of standard procedures).
Task design	Training in the design of multiple choice questions (MCQs) and short answer questions (SAQs). Design of 'stepped' or structured questions. Estimating student workload associated with tasks; designing optimum teacher assessment load. Design to constrain plagiarism and malpractice.
Fair practice	Designing and appraising tasks so that they are accessible to people with disabilities. Designing and reviewing tasks to ensure that they are accessible regardless of race, age, religion, gender.
Forms of judgement	Recognising three forms of judgement (scientific, legal and aesthetic). Appreciating the appropriateness of each form to different sorts of learning outcomes and recognising the differences in the ways in which each form describes achievement.
Granularity	Understanding that judgements can be made simply (as in the low-granularity pass-fail) or with high granularity (marking on a 0-100 scale). Knowing the costs and technical issues associated with reliable judgement at different levels of granularity.
Improving reliability and reducing uncertainty	Knowing how to improve marker reliability by means of a manageable number of usable criteria. Knowing how to write usable and clear criteria. Understanding that reliability is enhanced by multiple judgements of an outcome, made by different judges, on the basis of a variety of tasks, over time.
e-assessment	Knowing how to create on-line tests to engage students, provide them with feedback for future learning and improve the reliability of judgements of the outcomes addressed by the tests.
Plagiarism	Knowing how to advise students on avoiding plagiarism. Knowing how to detect plagiarism. Considering assessment designs that are inimical to plagiarism.
Recording achievement	Knowledge of ways of recording and communicating achievements that are not described by numbers. Portfolios, records of achievement and claims making.
Course assessment plans	Creating a public document that clearly describes what will be assessed, how, why and when. Rules, regulations, classification and weighting systems should be included. Ideally, links will be made with programme assessment plans.

These recommendations are limited because they assume that it is sufficient to be thoughtful about one's own course or module in isolation from the degree programme as a whole. While this may be a sensible coping strategy and attract little, if any, criticism, it is not ideal because

students experience, and universities provide, programmes of study that lead to awards – and awards often warrant competence achieved through years of study. Certainly, there is concern in the pbpl Centre that programme-level assessment arrangements are not sufficiently robust to capture complex learning outcomes. The point is developed in Appendix 2.

There is a case to be developed (and, preferably, to be tested by empirical enquiries) that problems which ought to attach to the assessment of 'wicked' competences are *programme-level* problems. In other words, they can be contained at the module or course level by means of coping or satisficing practices. In terms of assessment theory, these practices may not be ideal but the assessment challenge really presents when it comes to warranting achievement at the end of a programme of study.

If that is plausible, the recommendations for enhancing individuals' assessment practices have value but are defective by failing to challenge the assumption that the best unit of analysis when it comes to thinking about student learning is the course/module. Arguably, it is the programme or award.

Departments

An enlightened department would be concerned with assessment at the programme or awards level. It would take a systemic or coherent approach. It might be concerned with most, if not all, of these themes and have some commitment to working them out in departmental policies and practices.

Topic	Design features [taking the programme as the unit of analysis]
Variety of assessment task	Programmes will use a good range of assessment tasks to maintain student engagement and address the full range of learning outcomes described in the programme specification (http://www.qaa.ac.uk/academicinfrastructure/programSpec/guidelines06.pdf).
Assessing the range of learning achievements	Programmes will be assured that the full set of learning outcomes is reached by assessment practices. They will also need to ensure continuity of attention, so that some outcomes are touched once and then ignored, and progression, so that level 3 performances are judged differently than level 1 performances on the same outcome.
Fair practice	There will be programme guidance on ensuring fair practice in assessment. There may be training to support it.
Forms of judgement	There will be programme guidance on the three forms of judgement (scientific, legal and aesthetic), their distinctive features, appropriateness and costs. Programme assessment plans will advise on appropriate uses and balance.
Granularity	Programme assessment plans will advise on the levels of granularity appropriate to different types of task and forms of judgement, possibly differentiating by level as well.
Improving reliability and reducing uncertainty	<p>The programme specification identifies programme learning outcomes. There will also be public descriptions, intelligible to students and, arguably, to employers as well, that say what counts as an acceptable level of achievement. Such criteria may be written for several levels of achievement – not adequate, pass and excellent, for example.</p> <p>Programme assessment maps will be used to establish that, where outcomes are to be warranted, then multiple judgements of each are made by different judges, on the basis of a variety of tasks, over time.</p>
e-assessment	Teams will provide training for staff and students in e-assessment and support its development throughout programmes.
Plagiarism	Programmes will have policies and guidance on plagiarism, which should be echoed in course handbooks. They will advise on detecting plagiarism and provide training in task designs that resist plagiarism.
Accumulation of marks and recording achievement	<p>There will be understanding of the implications of different rules, regulation and accumulation practices on student progress and degree classification. Assessment databases and information management systems will be in place.</p> <p>There will be means to record data that are not in numerical form.</p>
Programme assessment plans	<p>There will be a public document that clearly describes what will be assessed, how, why and when. Student assessment entitlements will be set out. A series of assessment maps shows (a) which assessment methods are used in the programme and where (b) where the programme learning outcomes are reached by assessment arrangement (c) how student assessment entitlements are located in the programme.</p> <p>Rules, regulations, classification and weighting systems are be included.</p>

Topic	Design features [taking the programme as the unit of analysis]
Communicating achievement	It will be understood that some sorts of statement about achievement can be reported with greater confidence than others
Assessment economics	Assessment has direct costs (the development of e-assessment systems, for instance), opportunity costs (time spent marking is not available for other purposes) and student costs (excessive or defective assessment practices can impeded learning and lower retention rates). Teams have considered these costs in the formulation of programme assessment plans and when appraising course assessment practices.
Educational development	The programme educational development plan enhances shared assessment capacity.
Assessment for transfer	There is a view that we have to work at transferring learning from one context to another. The further the psychological distance between the two contexts, the less likely it is that there will be transfer and the more effort is needed to encourage it. Programmes that are serious about encouraging the transfer of learning need to design assessment tasks that support 'far' transfer as well as the more common 'near' transfer.
Accreditation of prior <i>experiential</i> learning (APEL)	With the increase, especial in professional areas, of process-centred and enquiry-based courses, there will be an advantage to programmes with robust arrangements for recognising prior, non-formal learning.
Promotion of an assessment-and-learning culture	It is sometimes assumed that students will pick up the rules of a programme's assessment game. They may pick up some rules but they may not be those associated with ambitious curricula. If assessment is to be more than a ritual in which students put in enough work on routine tasks to get acceptable grades, then hard work needs to be done to promote the new assessment-and-learning culture.

It is an open question whether departments are sufficiently coherent and whether departmental heads have adequate preparation in curricular and pedagogic leadership to meet these expectations. The quality of departmental leadership has been a concern in the USA and UK for some years (Wergin, 1994) as has curriculum coherence (Gaff and Radcliffe, 1996); there is no compelling evidence that concerns have been assuaged.

Institutions

Universities provide policies and regulations that delineate preferred learning-and-assessment cultures and inform programme or departmental designs. Enlightened universities might have the following concerns in the design and management of assessment practices.

Topic	University actions
Variety of assessment task	Universities will usually have assessment policies that delineate procedural requirements (examinations, accessibility, classification processes and regulations). They will also set out some academic expectations, saying that students are entitled to sufficient variety of tasks to assess the full range of programme learning outcomes. This might be taken as a statement of the process standards that programmes should meet.
Assessing the range of learning achievements	It is helpful to comment on the balance between tasks with formative and summative purposes. Modules will generally contain both types; some learning outcomes are more likely to be judged for formative purposes than for summative; maximising reliability is a big concern with summative tasks.
Fair practice	The University recognises that transcripts and certificates are easier to judge when accompanied by statements about output standards and reliability.
Differentiated assessment	Since there is a trend towards 'blended learning' and a strong possibility that 'Web 2.0' will transform higher education models, it is likely that university policy will encourage the development of appropriate and efficient e-assessment arrangements.
Improving reliability and reducing uncertainty e-assessment	University policies and regulations will govern programme and departmental policies.
Plagiarism	University policies and regulations will govern programme and departmental policies.
Cumulation of marks and recording achievement	Universities will know that their database systems affect the ways in which marks are transformed and cumulated.
Programme assessment plans	Universities often set out expectations of programme assessment plans.
Communicating achievement	Warrants may be accompanied by explanatory text, transcripts and portfolios (or personal development plans). University expectations need to be delineated.
Assessment economics	Information about process standards and other external indicators of programme quality would help outsiders interpret local warrants.
Educational development	Policies will remind programme teams of the costs of assessment. They may refer teams to sources of advice on efficient assessment.
Assessment for transfer	Educational development support should be provided for departments, teams and individuals.
Accreditation of prior <i>experiential</i> learning (APEL)	Policies will remind programme teams of the value of setting 'far' transfer as well as 'near' transfer tasks.
Promotion of an assessment-and-learning culture	Universities are likely to encourage the adoption of APEL and delineate some standard approaches. Web 2.0 developments and emphasis on practice-based learning make this an important aspect of universities' business model.
	Policies are often accompanied by a tolerably-accessible explanation of the evidence and other thinking behind them. The idea that assessment-and-learning cultures need to be created and systematically maintained may be novel enough to need such attention.

Students

Work on enhancing student employability (for example, Knight and Yorke, 2004) has argued that students:

- Should have assessment entitlements – should be entitled to a set of assessment experiences and practices.
- Should be 'knowing' students – should understand assessment purposes, practices, criteria, entitlements and ways of representing their learning to others.
- Should be participants in assessment – especially by peer- and self-assessment. In the case of practice-based learning, they will often be negotiating assessment tasks and criteria as well.

This work argues that such practices are particularly appropriate when it comes to assessing 'complex' achievements, which may be taken as tantamount to 'wicked' competences. Broadly speaking, this is the approach taken by Alverno College, Milwaukee, which has established an international reputation for its assessment, learning and teaching nexus. An interesting point is that the competences fostered by all Alverno programmes can, to a greater or lesser extent, be described as complex or 'wicked'. They are:

- Communication
- Analysis
- Problem solving
- Valuing in decision-making
- Social interaction
- Developing a global perspective
- Effective citizenship
- Aesthetic responsiveness

Alverno routinely engages students in the assessment process, expecting them to understand and apply criteria through peer- and self-assessment, as well as negotiating tasks and outcomes in complex, authentic practice contexts.

This fits with recent work by Boud and Falchikov (2006), who argue that it assessment practices ought to encourage students to set out as lifelong learners. This involves helping them to become assessors themselves, making them participants in assessment practice, rather than its objects. A problem is that this radical approach is at odds with traditional notions, which may have 'been so contaminated and associated with actions that students wish to avoid that the notion of becoming a lifelong assessor is anathema'. (p. 407). There follow ten illustrations of ways of thinking about teaching, learning and assessment practices that are aligned with an emphasis on life-long learning. Such practice:

- Engages with standards and criteria and problem analysis
- Emphasises importance of context
- Involves working in association with others
- Involves authentic representations and productions
- Promotes transparency of knowledge
- Fosters reflexivity
- Builds learner agency and constructs active learners
- Considers risk and confidence of judgement
- Promotes seeking appropriate feedback
- Requires portrayal of outcomes for different purposes. (pp. 408-410)

There appears to be a good fit between this view of teaching, learning and assessment and the concerns of the pbpl Centre. In that spirit, we suggest that, especially where the assessment of 'wicked' competences is concerned, assessment practices engage students as collaborators. This would entail movement away from the predominant 'measurement' model of assessment, at

least insofar as 'wicked' competences are concerned. Following the commentary in pages 3-11, this would be a wise trajectory.

It should also help students to represent better their learning processes and achievements to employers, graduate schools and others.

References

- Bar-On, R. (1997). *Development of the Bar-On EQ-i: A measure of emotional intelligence and social intelligence*. Toronto: Multi-Health Systems.
- Barrie, S. C. (2004) A research-based approach to generic graduate attributes policy, *Higher Education Research and Development*, 23(3), 261-275.
- Belbin, R. M. (2003) *Management Teams – Why They Succeed or Fail*. London: Butterworth Heinemann.
- Bostrom, R. French, R. Johnson-Laird, P. Parshall, C. (2004). *Review and Evaluation of the Development of a Listening Comprehension (LC) Section of the LSAT*. Newtown, Pennsylvania, Law School Assessment Council. (<http://www.uky.edu/~bostrom/listfinal!.htm>, accessed 30 August 2006)
- Boud, D. and Falchikov, N. (2006) Aligning assessment with long-term learning, *Assessment and Evaluation in Higher Education*, 31(4), 399-413.
- Bradberry, T. and Greaves, J. (2005). *The Emotional Intelligence Quickbook*, New York: Simon and Schuster.
- Brennan, J., Johnstone, B., Little, B. Shah, T and Woodley, A. (2001) *The Employment of UK Graduates: comparisons with Europe and Japan*. London: The Higher Education Funding Council for England.
- Brown, J. S. and Duguid, P. (2000) *The Social Life of Information*. Cambridge MA: Harvard University Press.
- Bryne, B. (1998) *Complexity Theory and Social Science*. London: Routledge.
- Carless, D., Joughin, G. & Mok, M. (2006) Learning-oriented assessment: principles and practice. *Assessment and Evaluation in Higher Education*, 31(4), 395-398.
- Checkland, P. (1981). *Systems Thinking, Systems Practice*. Chichester: John Wiley.
- Claxton, G. (1998), *Hare Brain, Tortoise Mind*. London: Fourth Estate.
- Colby, A., Ehrlich, T., Beaumont, E. and Stephens, J. (2003) *Educating Citizens. Preparing America's Undergraduates for Lives of Moral and Civic Responsibility*. San Francisco: Jossey-Bass
- Conklin, J. (2003) Dialog Mapping: An Approach for Wicked Problems, *CogNexus Institute*, <http://cognexus.org/dmforwp2.pdf>, accessed 25 August 2006.
- Dreyfus, H. and Dreyfus, S. (2005) Expertise in real world contexts, *Organization Studies*, 26(5), 779-792.
- Dweck, C. (1999) *Self-theories: their role in motivation, personality and development*. Philadelphia: Psychology Press.
- Eysenck, H. (2000). *Intelligence: A New Look*, Transaction Publishers.
- Feltovich, P. J., Hoffman, R. R., Woods D. and Roesler, A. (2004) Keeping it simple: how the reductive tendency affects cognitive engineering, *IEEE Intelligent Systems*, May/Jun, 90-94.
- Gaff, J. G. and Ratcliff, J. L. (Eds.) (1996) *Handbook of the Undergraduate Curriculum*. San Francisco: Jossey-Bass.
- Gardner, H. (1983) *Frames of Mind*, New York: Basic Books.
- Gardner, H. (2007) *Five Minds for the Future*. Cambridge MA: Harvard Business School Press.
- Goleman, D. (1996). *Emotional Intelligence: why it can matter more than IQ*. London: Bloomsbury.
- Harvey, L., Moon, S. and Geall, V with Bower, R., 1997, *Graduates' Work: Organisation change and students' attributes*. Birmingham, Centre for Research into Quality (CRQ) and Association of Graduate Recruiters (AGR).
- Huber, M., Brown, C., Hutchings, P., Gale, R., Miller, R. and Breen, M. (Eds.) (2007) *Integrative Learning: Opportunities to Connect. Public Report of the Integrative Learning Project*. (Stanford, CA: the Association of American Colleges and Universities and The Carnegie Foundation for the Advancement of Teaching).
- <http://www.carnegiefoundation.org/files/elibrary/integrativelearning/index.htm>, accessed 30 January 2007
- Jacques, D. (2000), *Learning in Groups: A Handbook for Improving Group Work*, Third Edition. London: Kogan Page.
- Jessup, G. (1991) *Outcomes: NVQs and the Emerging Model of Education and Training* (London: Falmer).

- Knight, P. T. (2001) *Small-scale Research: pragmatic enquiry in social science and the caring professions*. London: Sage
- Knight, P. T. (2002) Summative assessment in higher education: practices in disarray, *Studies in Higher Education*, 27(3), 275-286.
- Knight, P. T. (Rapporteur) (2003) *Measures to improve HE/R in order to strengthen the strategic basis of the ERA. The report of an independent high level expert group set up by the European Commission*. Luxembourg: Office for Official Publications of the European Community.
- Knight, P. T. and Yorke, M. (2006a) *Assessment close up: the limits of exquisite descriptions of achievement*. Paper presented at the Higher Education Close-up conference, Lancaster, 26th July.
- Knight, P.T. (2006b) Assessment as local practice, *Assessment and Evaluation in Higher Education*, 31(3), 435-452
- Knight, P. T. and Yorke, M. (2004) *Learning, Curriculum and Employability in Higher Education*. London: Routledge.
- Law, J. (2000) *Objects, Spaces and Others*, the Centre for Science Studies, Lancaster University, Lancaster LA1 4YN, UK, at <http://www.comp.lancs.ac.uk/sociology/papers/Law-Objects-Spaces-Others.pdf>
- Law, J. (2004) *After Method: Mess in Social Science Research*. London: Routledge.
- Loacker, G. (2005) *Enhancing Assessment through Teaching and Learning*. Paper presented to the first international conference on teaching and learning through assessment, Hong Kong, 14th July.
- Maki, P. L. (2004) *Assessing for Learning Building a Sustainable Commitment Across the Institution*. Sterling VA: Stylus Publishing.
- Mason, G. Williams, G., Cranmer, S. Guile, D. (2003) *How Much Does Higher Education Enhance the Employability of Graduates?* Bristol: the Higher Education Funding Council for England.
- Mayer, J.D. & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, 17, 433-442.
- Norcini, J. (2005) Current perspectives in assessment: the assessment of performance at work, *Medical Education*, 39, 880-889.
- Nursing and Midwifery Council's (2005?) *Standards of Proficiency for Pre-registration Nursing Education*. London: The UK Nursing and Midwifery Council.
- The Nursing and Midwifery Council (2006) *Standards to support learning and assessment in practice: NMC standards for mentors, practice teachers and teachers*. London: Nursing and Midwifery Council.
- Oppenheim, A. N. (1992) *Questionnaire Design, Interviewing and Attitude Measurement*. London: Pinter.
- Perkins D (2006) Constructivism and troublesome knowledge in Meyer JHF and Land R 2006 (Eds) *Overcoming Barriers to Student Understanding: Threshold Concepts and Troublesome Knowledge*. Abingdon, Oxford: RoutledgeFalmer
- Price, M. and Rust, C. (1999) The experience of introducing a common criteria assessment grid across an academic department, *Quality in Higher Education*, 5(2) 133-144
- Qualifications and Curriculum Authority (2004) *Foundation stage: 2003/4 annual report*. London: Qualifications and Curriculum Authority.
- Quality Assurance Agency for Higher Education (2000a) *Honours degree benchmark statements: Accounting*. Gloucester: The Quality Assurance Agency for Higher Education.
- Quality Assurance Agency for Higher Education (2000b) *Honours degree benchmark statements: Social Policy and Administration and Social Work*. Gloucester: The Quality Assurance Agency for Higher Education.
- Quality Assurance Agency for Higher Education (2006) *Outcomes from institutional audit Assessment of students*. Gloucester: The Quality Assurance Agency for Higher Education.
- Rhodes, S. C., Watson, K. W. and Barker, L. L. (1990) Listening Assessment: Trends and influencing factors in the 1980s, *Journal of the International Listening Association*, 4, 62-82.
- Banta, T. et al., (Eds.) (1996) *Assessment in Practice* (San Francisco: Jossey-Bass)
- Rittel, H., and M. Webber (1973) Dilemmas in a General Theory of Planning, *Policy Sciences*, 4, pp 155-169

- Salovey, P. and Mayer, J.D. (1990). Emotional intelligence, *Imagination, Cognition, and Personality*, 9(1990), 185-211.
- Shay, S. (2005) The assessment of complex tasks: a double reading, *Studies in Higher Education*, 30 (6) 663-679.
- Shulman, L. (1986) Those who understand: knowledge growth in teaching, *Educational Researcher*, 15, 4-14.
- Sternberg, R. J. and Grigorenko, E. L. (2000) Practical intelligence and its development, in R. Bar-On, and J. Parker (Eds.) (2000) *The Handbook of Emotional Intelligence*. San Francisco: Jossey-Bass, 215-243.
- Swanwick, T. (2005) Informal learning in postgraduate medical education: from cognitivism to 'culturism', *Medical Education*, 39: 859-865.
- Tett, R. P., Fox, K. E., & Wang, A. (2005). Development and validation of a self-report measure of emotional intelligence as a multidimensional trait domain. *Personality and Social Psychology Bulletin*, 31, 859-888.
- The General Teaching Council (2006) *Statement of Professional Values and Practice for Teachers* (London: The General Teaching Council, <http://www.gtce.org.uk/standards/disc/StatementOfProfValues>, accessed 25 August 2006),
- The National Youth Agency (2004) *The NYA Guide to Youth Work and Youth Services*: Leicester: The National Youth Agency.
- Thrift, N. (1999) The place of complexity, *Theory, Culture and Society*, 16(3), 31-69.
- Training and Development Agency for Schools (2006) *Qualifying to teach: Professional standards for qualified teacher status and requirements for initial teacher training*. London: Training and Development Agency for Schools.
- van Geert, P. (1994) *Dynamic Systems of Development: change between complexity and chaos*. Hemel Hempstead: Harvester Wheatsheaf.
- Walvoord, B. E. and Anderson, V. J. (1998) *Effective Grading: A tool for learning and assessment*. San Francisco: Jossey-Bass.
- Wergin, J (1994) *The Collaborative Department*. Washington DC: The American Association for Higher Education.
- Yorke, M. (1999) The skills of graduates: a small enterprise perspective, in: O'Reilly, D. Cunningham, L. and Lester, S. (eds.) *Developing the Capable Practitioner*. London: Kogan Page, 174-183.
- Yorke, M. (2005a) *Issues in the Assessment of Practice-based professional Learning*. Milton Keynes, The Open University, <http://www.open.ac.uk/cetl/pbpl/links/MYNov05.pdf>, accessed 25 August 2006)
- Yorke, M. (2005b) Formative assessment and employability: some implications for higher education practices, in: McNay, I. (Ed.) *beyond Mass Higher Education*. Maidenhead, The Open University Press and Society for Research into Higher Education.

Appendix 1: Extracts from the Wikipedia entry 'wicked problem'.

(http://en.wikipedia.org/wiki/Wicked_problems, accessed August 25 2006)

The concept of "wicked problems" was originally proposed by H. J. Rittel (a pioneering theorist of design and planning, and late professor at the University of California, Berkeley) and M. Webber in a seminal treatise for social planning. Rittel expounded on the nature of ill-defined design and planning problems which he termed "wicked" (i.e.. difficult) to contrast against the relatively "tame" problems of mathematics, chess, or puzzle solving. Wicked problems have incomplete, contradictory, and changing requirements; and solutions to them are often difficult to recognize as such because of complex interdependencies. Rittel and Webber stated that while attempting to solve a wicked problem, the solution of one of its aspects may reveal or create another, even more complex problem.

Classic examples of wicked problems include economic, environmental, and political issues (for an extreme case, consider what it would take to "solve" terrorism, where even the term terrorism is highly controversial and difficult to define).

...

According to Conklin (2003), the four defining characteristics of wicked problems are:

1. The problem is not understood until after formulation of a solution
2. Stakeholders have radically different world views and different frames for understanding the problem.
3. Constraints and resources to solve the problem change over time.
4. The problem is never solved.

The following characteristics further describe wicked problems:

- Wicked problems do not have an exhaustive set of potential solutions.
- Every wicked problem can be considered to be a symptom of another problem.
- Discrepancies in representing a wicked problem can be explained in numerous ways -- the choice of explanation in turn determines the nature of the problem's resolution.
- Every wicked problem is essentially unique -- lessons-learned are hard to transfer across to other problems.
- Wicked problems are often "solved" (as well as they can be) through group efforts.
- Wicked problems require inventive/creative solutions.
- Every implemented solution to a wicked problem has consequences, and may cause additional problems.
- Wicked problems have no stopping rule(s).
- Solutions to wicked problems are not true-or-false, but instead better, worse, or good enough.
- There is no immediate and no ultimate test of a solution to a wicked problem.
- The planner or designer (solving the problem) has no inherent right to solve the problem, and no permission to make mistakes.

Appendix 2. Extracts from pbpl document *Developing systems to support pbl*, January 2007

Many of our programmes of study are designed to provide a basis for the formation of professional competence or expertise. This often implies the kind of complex learning outcomes which need to develop across a programme of study rather than within a single course. At the same time the award rather than course orientation of external accreditors tends to emphasise and require evidence for award level outcomes. At the same time the need to design assessment which is capable of evidencing practice skills and competences often leads to expensive forms of assessment.

However both OU assessment policies and systems require assessment to be course based and that every course (except level 1) has an examinable component. In consequence, assessment is both more expensive than necessary across our awards and often not sufficiently well configured to deliver and measure award level learning outcomes. If we had the option for some awards to select the optimal amount of examinable components across an award, to allow some assessment process (e.g. portfolios) to build across multiple courses, and to replace course level exam boards with a single award board we could both reduce costs and improve assessment quality. The opportunity to reduce the burden of (summative) assessment on students may also reduce student fatigue and drop out and make room for more effective learning.

A further issue is that assessment systems assume that continuous assessment components will be submitted in the form of a TMA. Thus ensuring the system captures and manages practice assessment in workplace settings can require the exercise of some ingenuity ...

... Professional qualifications aim to extend and enhance students' understanding and aptitudes. However, many students find that they already have mastery of outcomes in the programme specification and (a) can be irked by having to go through concepts and material with which they are already perfectly familiar (b) would consider applying for a programme if existing competence and achievement could be recognised.

Both The pbpl CETL and COBE have been working on APEL (assessment of prior experiential learning) processes that could meet these circumstances. We seek to develop a university framework for APEL.